BIRKHÄUSER

Felix Klein
Arnold Sommerfeld

The Theory of the Top
Volume I

Introduction to the Kinematics and Kinetics of the Top

Raymond J. Nagem
Guido Sandri
Translators

Preface by Michael Eckert

Birkhäuser
Boston • Basel • Berlin

Raymond J. Nagem
Boston University
Boston, MA 02215
USA
nagem@bu.edu

Guido Sandri
Boston University
Boston, MA 02215
USA
sandri@bu.edu

ISBN: 978-0-8176-4720-9 e-ISBN: 978-0-8176-4721-6
DOI: 10.1007/978-0-8176-4721-6

Library of Congress Control Number: 2008926224

Mathematics Subject Classification (2000): 01A75, 70E05, 70E15, 70E17, 70E50

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com

Contents

Chapter III
The Euler equations, with further development of the kinetics of the top

Preface

The *Theory of the Top* attained its great fame from both its monumental scope and its outstanding authors. In the early twentieth century, Felix Klein was known as a mathematician of world fame; Arnold Sommerfeld, Klein's disciple, had acquired his reputation as a rising star of theoretical physics. By 1910, when the final volume of this treatise was published, the names of Klein and Sommerfeld would signal to a student that a matter as complex as the top was presented in a most authoritative manner, from the perspective of both mathematics and physics. The work also stands out in other regards: by its sheer extent—four volumes comprising a total of almost a thousand pages—and by the time lag of about fifteen years between inception and completion. Klein himself regarded the final result as somewhat disjointed. Its "idiosyncratic disposition," he reflected in 1922, may be understood only by taking into account the historic circumstances at its inception in 1895; the developments between the first and last parts derailed the project from its intended course, so that for the technical applications described in Volume IV "almost no use was made of the theoretical framework developed at the beginning" [Klein 1922, p. 659].

It seems appropriate, therefore, to recall the historical circumstances under which this treatise was conceived and pursued. Felix Klein was not only a renowned mathematician, but also an entrepreneurial and ambitious university professor striving for a broader acknowledgment of mathematics as a cultural asset. During the Wilhelmian Era, when Germany was struggling for recognition as a great power, cultural affairs were no longer innocent bystanders of national politics. Friedrich Althoff, a powerful reformer at the Prussian Ministry of Culture, attempted to form centers of excellence at certain universities [Brocke 1980]. Klein had become professor in Göttingen University in 1886.

After a few years of frustration and uncertain prospects, Klein persuaded Althoff that Göttingen would assume the desired rank only if rising stars like David Hilbert and Hermann Minkowski were called to the university as his colleagues. As a result of Klein's strategy, backed by the almighty Althoff, Göttingen became a mecca of mathematics [Rowe 1989].

But Klein's ambitions were not restricted to local affairs at his university. As a part of his attempts to gain widespread recognition for mathematics, he began to edit an *Encyclopedia of Mathematical Sciences*, an enterprise that lasted until the 1920s and encompassed, in addition to pure mathematics, a broad spectrum of mathematical applications to mechanics, physics, and astronomy. Klein also established contacts with the Association for the Advancement of Mathematical and Scientific Education, in order to gain influence on high school teaching of mathematics. Furthermore, he displayed considerable interest in the scientific training of engineers, which was traditionally the realm of technical universities, and therefore made Klein the enemy of engineering professors who regarded his tendencies as an unwelcome interference in their own affairs. In 1895, for example, Klein conceived a memorandum in which he suggested the foundation of a new institute in Göttingen University for the education of the "general staff" of technology, whereas the training of "front officers" could be left to the technical universities [Rowe 1989, p. 203].

Such was the broader context for the birth of the *Theory of the Top.* Under this title Klein announced a special lecture in the winter semester of 1895/96, addressed to high school teachers who wished to keep in touch with advanced mathematical subjects. In the preceding semester, Klein had held another special lecture for the same audience on *Elementary Geometry.* One of his assistants was charged with elaborating the manuscript of this lecture into a booklet, which Klein presented to the high school teachers association as a special gift by which he intended to prepare the ground for his further engagement in aiming at a general reform of high school teaching. By lecturing on the top in the winter of 1895/96, Klein attempted to demonstrate that his university teaching was not an ivory tower activity but had relations to technological as well as educational affairs. Like the *Elementary Geometry* of the preceding semester, the *Theory of the Top* was meant to be printed afterward as a small booklet and presented as a gift to his extramural clients. Klein remarked in an autobiographical note in 1913

that the *Top* was a tactic intended as a second dedicatory publication [Jacobs 1977, p. 18].

The course of subsequent events, however, precluded a smooth realization of these plans. Klein entrusted Arnold Sommerfeld, who had become his assistant in the autumn of 1894, with much more than the elaboration of this lecture. Sommerfeld had just accomplished his habilitation (the German ritual to acquire the right to lecture in a university) on the theory of diffraction [Sommerfeld 1896] and was busy with the elaboration of Klein's *Number Theory*, a lecture that Klein held in two parts in the winter of 1895/96 and summer of 1896 [Klein 1896a; Klein 1896b]. When Sommerfeld finally started to work on the theory of the top in the autumn of 1896, he did so without great enthusiasm. He worked on several projects at the same time, all of them related to one or another of Klein's activities, such as a register for the *Mathematische Annalen* (a journal edited by Klein) or a review article on partial differential equations for Klein's *Encyclopedia of Mathematical Sciences*. Furthermore, he discovered that the methods developed in his habilitation work proved to be more fertile than he had originally anticipated. Writing papers on his own research appeared more interesting to him than editing Klein's lecture on the theory of the top.

Although the first parts of the lecture advanced to the state of proof reading by the spring of 1897, its completion was dragging on. In March of 1897 Sommerfeld wrote to Klein that "the number of boundary value problems that I am able to solve by my extension of Thomson's mirror method is very considerable." He felt sure that Klein would appreciate the temporary neglect of the top, because his method for solving physical differential equations was completely in line with Klein's tendencies: "I hope you will enjoy it yourself. But I still have several days to do with it. If you could arrange for this work to be published soon in an English journal, such as the London Math. Soc. [*Proceedings of the London Mathematical Society*], I would be very happy." To please Klein he added some remarks about his elaboration of the theory of the top, but finally revealed that this had a rather low priority on his to-do list. "Unfortunately, I have to admit that in the meantime the top has been in the nonetheless very interesting 'sleeping top' state" [Sommerfeld 1897a].

Working under Klein must have been quite demanding. "I really cannot write to you each day," Sommerfeld once apologized to his fiancée Johanna Höpfner. "Klein's bullwhip is rather close behind me" [Sommerfeld 1897b]. At some point in 1897, Klein must have decided

to split the publication into several parts. Klein did not leave the theory of the top in the state in which he had presented it in his lecture of 1895/96. In the summer semester of 1896 he lectured on technical mechanics. In October and November of 1896, he chose the theory of the top as a theme for guest lectures at Princeton University [Klein 1897]. In view of Klein's goal of demonstrating the uses of mathematics to engineers at technical universities, he must have regarded it expedient to include more applied matters and charged Sommerfeld to work out the details.

Under these premises, the mathematical foundations as laid out in earlier lectures were published in 1897 as Volume I. As Sommerfeld prepared the subsequent volume, progress became slow because he struggled with problems that were the subject of controversial debates. "With regard to the equilibrium stability H.[adamard] does not go one step further than Lyapunov"; such remarks in the correspondence between Sommerfeld and Klein [Sommerfeld 1898a] illustrate that subjects like stability, dealt with in a chapter of the second volume, could easily give rise to new debates and prevent rapid publication. Nevertheless, Sommerfeld completed the second volume without much delay, so that it appeared just a year after the first volume in 1898. As the first reactions made evident, the more subtle parts of the book such as the chapter on stability provoked criticism: "I would have a number of remarks about your definition of stability," Heinrich Burkhardt commented after the appearance of Volume II. "But I would need a day or two to formulate them clearly and precisely, which I do not now have. It seems to me that in your definition stability is the rule, instability the exception ... I tend to guess that all motions of the top are stable according to this definition, except those whose instability you have proven" [Burkhardt 1898].

Such reactions cautioned against rushing to publication—all the more because the plan for the remaining parts addressed subjects beyond the realm of mathematics proper: gyroscopic phenomena in geophysics, astronomy, and technology. Employing mathematical virtuosity in these fields was easier proclaimed than done. But Sommerfeld was not afraid to meet this challenge. For example, he corresponded extensively with the naval engineer Carl Diegel in Kiel, the German naval base, about the application of the theory to the gyroscopic guidance of torpedos [Diegel 1898]. In December of 1898 he wrote to Klein that "My letters to D.[iegel] have the extent of treatises. I may travel from Göttingen back to Cl.[austhal] via Kiel. In any case

this correspondence gives rise to a nice paragraph about 'applications of theory in technology' " [Sommerfeld 1898b].

Since the autumn of 1897, Sommerfeld had been professor of mathematics at the mining academy in Clausthal, so that his communication with Klein had to occur via the exchange of letters and at occasional meetings, which further retarded the project. In addition, Klein persuaded Sommerfeld in 1898 to join his *Encyclopedia* project as an editor for the planned volumes on physics, a capacity that contributed to the derailment of Sommerfeld's career as a mathematician; he transformed more and more from a mathematician into a theoretical physicist, a metamorphosis that is also reflected by the choice of his research papers at the turn of the century. "Unfortunately I had no time for the top," he apologized in a letter to Klein in November 1899. "I have to get rid of my [paper on] X-rays before I can deal with something else" [Sommerfeld 1899]. Klein responded that he was "thinking of our top with silent sorrow" [Klein 1899]. In 1900, Sommerfeld exchanged his position as a professor of mathematics in Clausthal for a professorship of mechanics at the technical university in Aachen. Although this brought him into closer contact with technological applications of gyroscopic theory, it did not accelerate the publication of the pending volumes. "When I will have time to resume the top?" he responded to Klein's urging in November 1900. "The entire next week there are examinations without interruption. I will hope for the best but promise nothing" [Sommerfeld 1900].

To cut a long story short, it took five years after the appearance of Volume II in 1898 before the third volume of the *Theory of the Top* was published, and seven more years before the fourth and final volume appeared in 1910. In the meantime, gyroscopic theory itself had advanced or was made the subject of other reviews. In 1907, for example, Klein admonished Sommerfeld to pay attention to a recent article of Paul Stäckel, who was writing on the top for the *Encyclopedia* [Klein 1907]. In the foreword to the fourth volume, dated April 1910, Klein and Sommerfeld had to admit that during this long time span "the unity of substance and manner of presentation was lost." The loss of unity and coherence was caused not just by a turn from mathematical foundations to technological applications. Despite Sommerfeld's close contacts with technology, his presentation of applied subjects in Volumes III and IV was written from the perspective of a mathematician and theoretical physicist, so that it did not really address engineering concerns. With regard to the technology of the gyrocompass, for

example, Sommerfeld admitted later that even in the part about the technical applications the text "nowhere addresses technical details" [Broelmann 2002, p. 138].

The *Theory of the Top*, therefore, is a strange monument of scientific literature from the turn of the nineteenth to the twentieth century: too heterogeneous to please one or another orientation, and yet outstanding in its scope and detail. Klein and Sommerfeld hoped that its versatility would be considered as a compensation for its "lack of systematicness." In the end, they confessed that the top was for them what it had been already for many natural philosophers in the nineteenth century: a target of opportunity for "awakening the sense for true mechanics," a "philosophical instrument."

Michael Eckert
Deutsches Museum, Munich

Translators' Remarks

The ordinary difficulties of a translation are happily moderated, in our present case, by the extraordinary greatness of the original work and its authors. We could do no better than to preserve the form and notation of the text in a literal and unabridged rendering. Our notes are added separately, and are intended primarily to provide historical context. They are indicated by numerical superscripts; the footnotes of Klein and Sommerfeld are retained, as are their own supplementary notes to Volume I, which were published as addenda when Volume IV appeared in 1910.

More comments in this place are not necessary. It remains for us only to acknowledge the very fine editors at Birkhäuser Boston, and to return, with pleasure, to the preparation of Volume II.

FOREWORD

When F. Klein gave a two-hour lecture "On the Top" in the winter semester of 1895/96, he attempted, in the first place, to emphasize the direct and particularly English conception of mechanical problems, as opposed to the more abstract coloring of the German school, and, on the other hand, to make the particularly German methods of Riemann function theory fruitful in mechanics. The consideration of applications and physical reality would thus be outlined and forcefully advanced in a detailed example, but not yet carried out to full extent.

In the extensive printed edition originating from the pen of A. Sommerfeld, interest in applications prevailed more and more, especially after his appointment to a teaching position in technical mechanics and later in physics. The astronomical, geophysical, and technical content added in this way required, in consequence, the necessity of a change, compared with the original lecture, in the mathematical point of view. While the approximation methods prepared in the first volumes (the method of small oscillations, the treatment of pseudo-regular precession) and the intuitive formulation of the principles of mechanics by means of the impulse concept were perfectly conformable to applications, the advanced function-theoretical methods, the exact representation of the motion by elliptic functions, etc., were later found to be dispensable. Thus, for example, the parameters α, β, γ, δ and their related quaternion quantities, whose geometric meaning was elaborated in the first volume and whose analytic importance was given special emphasis in the second volume, withdrew in the third and fourth volumes, naturally in complete agreement with Klein himself, whose interests had likewise turned more and more toward applications. In particular, the presentation of the technical top problems in the fourth volume used only the very simplest and most elementary law of top motion, which flows immediately from the concept of the impulse in the dynamics of rigid bodies, and which is briefly derived once more at the beginning of this volume.

We would not deny, that with the loss of the unity of time in the course of the fifteen years which have elapsed between the first plan and the present conclusion of the book, our work has also lost its unity of substance and manner of presentation; that what we often promised

earlier with respect to the general, so-called analytic mechanics, especially in the advertisements of Vols. I and II, was later not kept; and that we have pursued many mathematical side roads which temporarily diverted us from our primary goal: the concrete understanding of dynamical problems. May the comprehensiveness of the content and the multiplicity of the engaged fields of interest be regarded as substitutes for the lack of systematicness and purposefulness of the presentation.

If we had to dispose of the collected subject material anew, we would probably present the actual mechanics of the top, including its applications, in a much smaller space, by pruning the analytic shoots that branch so joyously from the stem of mechanics. With this presentation we would address the large audience with scientific or technical interests in the theory of the top. The detailed analytic developments, which we would certainly not suppress simply on the basis of their special beauty, would be submitted in another presentation only to the more restricted mathematical circle. As for what pertains, finally, to the requirements of the completely unmathematical reader, and therefore to the difficult question of the popular explanation of the top phenomena, we have taken an extensively grounded critical position in the second volume, and at the beginning of the fourth volume have again pointed out the somewhat long but, it appears to us, only passable way that begins from the general impulse theorems of rigid body dynamics. The impulse theorems are either systematically developed from particle mechanics, or, should the occasion arise, illustrated only by experiments, and then postulated axiomatically; on the basis of these theorems, all the partly paradoxical facts of the theory of the top may be understood qualitatively as well-defined approximations, and their domains of validity delimited without want of clarity.

The top is suitable above all other mechanical devices for awakening the sense for true mechanics. May it, in the presentation of our book, serve this purpose in elevated measure, and thus prove worthy in the future of the honorable surname formerly bestowed upon it by Sir John Herschel,[1] the name of a philosophical instrument!

Göttingen and München, April 1910.

F. KLEIN. A. SOMMERFELD.

Advertisement of the Book

(from the notices of B. G. Teubner publishing company in Leipzig).

The work owes its origin to a lecture given by Prof. Klein during the winter semester of 1895/96 in Göttingen University. The elaboration of the ideas set forth in the lecture and the rounding out of the subject matter have since been the primary responsibility of Dr. Sommerfeld.

The first part, which appears in July of this year, presents, after a preparatory chapter of kinematic content, the fundamental considerations on the principles of mechanics, in so far as they apply to the present topic. A singular character of this section is the authors' frequent return, in the spirit of the older writers, to impact forces, and, throughout, to the concept of the "impulse" (W. Thomson's terminology; Poinsot's couple d'impulsion); that is, the impact turning-force that is able to produce the actual motion instantaneously from rest. The theory of the top, and the mechanics of rigid bodies in general, thus acquire a higher degree of clarity and simplicity than that obtained by the exclusive use of continuously applied forces.

The second part treats in detail of the mathematical side of the theory, the explicit representation of the motion of the heavy top by means of elliptic functions. It is shown here that neither the commonly used asymmetric Euler angles nor the symmetric Euler parameters (quaternion quantities), but rather certain parameters arising from Riemann function theory, are the simplest building stones, in analytic respects, from which the general formulas for the motion of the top may be composed.

The third part contains, in addition to many supplements to the previous material (consideration of friction at the support point, criticism of the popular top literature, etc.), the manifold applications of the theory to astronomical and physical questions. The accumulated treasures of the English literature, and especially the Natural Philosophy of Thomson and Tait, are of particular value here in presenting the investigations of

cyclic systems, gyrostats, etc., to the German public in a convenient readable form.

Originally conceived as a dedication to the Association for the Advancement of Mathematical and Scientific Education,[2] the book should also be understandable without difficulty to the more advanced research mathematician and physicist. Specific prior knowledge of analytic mechanics or function theory is not assumed. It is hoped, however, that the specialized mathematical circle will thus feel, without displeasure, a certain breadth and comfort in the presentation.

The tendency of the book may be characterized, finally, by a few sentences taken from the Introduction:

"The development of theoretical mechanics has taken, especially in Germany, an overly exclusive direction toward abstraction and formulas, which often detracts from a direct understanding. The student who learns well to derive the general principles of mechanics analytically does not always grasp their true mechanical meaning in a sufficiently lively sense, and often appears awkward when faced with obtaining the solution of a specific problem.

"We wish to oppose this recent and rising evil by a thorough treatment of our problem. We wish to establish not only a knowledge of mechanics, but also, so to speak, a feeling for it. Full clarity in the geometric aspects of motion is naturally a first prerequisite for this. ... Still more important for us, however, is full clarity concerning the forces that come into play as the mechanical causes of the motion. We will convey these forces as concretely as possible in space by vectors; we place special value on the development and consistent use of the impulse principle, etc. ... We do not intend, however, to minimize in any way the analytical side of our problem. The formula is ultimately the simplest and most concise description of the process of motion; it is indispensable, moreover, as the basis of actual numerical calculation. We will only demand, that our knowledge of mechanics be based not on formulas, but rather, on the contrary, that the analytic formulation appear of itself as the last consequence of a fundamental understanding of the mechanical principles."

Chapter I

The kinematics of the top

§1. Geometric treatment of the kinematics

We begin with a chapter of geometric content that treats of the *kinematics* of the top. In opposition to kinematics, we use, after the suggestion of Thomson and Tait, the word *kinetics*.[15] While kinematics operates merely with space and time and investigates motions only according to their geometric possibility, kinetics adds the concepts of mass and force, and treats of motions with regard to their mechanical possibility.

Among the properties of the top postulated in the Introduction, only the rigidity of the material and the fixed position of the support point come into consideration here, since the mass distribution of the body is completely irrelevant in kinematics. The following investigations are therefore valid for an *arbitrary rigid body with a fixed support point. We wish to denote such a body as a "generalized top," in contrast to the "symmetric top" defined in the Introduction.* In the immediately following chapters we also refer, on occasion, to this "generalized top," while in the later chapters we must limit ourselves entirely to the "symmetric top."

From the most general point of view, problems of kinematics are classified according to the *number of degrees of freedom.* The meaning of this expression, also introduced by Thomson and Tait, will be illustrated through the following small tabulation.

A freely moving point in space has *three degrees of freedom* (its position is determined by three independent coordinates).

A freely moving rigid body in space has *six degrees of freedom* (the position and orientation of the body are fixed through the specification of six appropriate independent parameters).

A rigid body with one point held fixed has again *three degrees of freedom.*

R.J. Nagem, G. Sandri, *The Theory of the Top*, DOI: 10.1007/978-0-8176-4721-6_I,

Our *top with a fixed point*, correspondingly, has *three degrees of freedom*, in so far as we may treat it as a rigid body. The *moving top* whose support point runs in a horizontal plane has *five degrees of freedom.* We can also construct a top with *one* or *two* degrees of freedom if we place the figure axis in a fixed frame or in a ring that rotates, in turn, about a fixed axis. On the other hand, our top has *infinitely many degrees of freedom* as soon as we wish to consider the *elastic deformation of the material.*

In the following considerations, we will begin with a freely moving rigid body in space, of which the top with a fixed support point is a special case. Since this is a very simple and well-known subject, it is enough to recall the relevant theorems briefly, without deriving them in detail. The proofs may be found, if necessary, in the previously cited textbooks.

We consider two different positions of a moving rigid body, and pose the problem of giving the motion that leads from the initial position to the final position in the simplest way. The position of the body is completely determined if the positions of any three of its points, say O, P, Q, are known. The initial positions of the points may be denoted by O_1, P_1, Q_1, and the final positions by O_2, P_2, Q_2. We can first transport the point O_1 to O_2 by a parallel displacement of the body; the points P_1, Q_1 are thus transformed into P_1', Q_1', respectively. We then connect P_1' with P_2 and Q_1' with Q_2, and construct at the center of the connecting lines the normal planes to the lines. These planes intersect in an axis that passes through O_2. We now rotate the body about this axis through an appropriate angle, so that P_1' is brought to P_2 and Q_1' is brought to Q_2. We can therefore transport the triad OPQ, and thus also the rigid body, from its initial position to its final position through a combination of a parallel displacement and a rotation.[16] Thus the theorem:

The most general change of position of a freely moving rigid body can always be replaced by a combination of a rotation and a translation.

If we take into consideration, further, that a parallel displacement is equivalent to a rotation about an infinitely distant axis, or a so-called rotation-pair (that is, two rotations about parallel axes with the same rotation angle but opposite sense[17]), then we can also give the previous

theorem in the following form, which is of interest with regard to the corresponding theorem for the statics of a rigid system:[18]

The most general change of position of a rigid body can be replaced by a single rotation and a rotation-pair.

Our construction can obviously be changed in a great variety of ways by replacing the initially chosen point O with some other point. We may designate the chosen point O as the "reference point," and can ask whether we can simplify the result of our construction through an appropriate choice of the reference point. In this respect, it results that one can always choose the reference point so that the direction of the translation and the axis of rotation are parallel. The combination of a rotation and a parallel displacement along the rotation axis is commonly called a *screw* (more precisely, a "motion-screw")[*]). The magnitude of the parallel displacement together with the magnitude of the rotation determine the pitch; the magnitude, axis, and sense of the rotation give the rotation angle, the rotation axis, and the rotation sense of the screw.[19] We can thus say:

The most general change of position of a rigid body can, by appropriate choice of the reference point, be replaced by a screw with a specific axis, a specific rotation angle and rotation sense, and a specific pitch.

Our screw-motion naturally coincides with the actual motion of the body only in the initial and final positions; the intermediate positions of the actual motion can be entirely different from the intermediate positions of the imagined screw-motion. Let us consider, however, an infinitesimal motion of the rigid body (that is, the limiting case of a finite motion during an infinitely diminished time interval) and the corresponding infinitesimal screw (that is, the limiting case of the corresponding finite screw-motion). Here we can no longer speak of intermediate states; consequently, we will regard an infinitesimal motion as directly identical to the constructed screw-motion and can state concisely:

Every infinitesimal motion of a rigid body is a screw-motion.

[*]) Cf. Sir Robert Ball. The theory of screws. Dublin 1876. (German edition by Gravelius, Berlin 1889).

In the following, we will characterize an infinitesimal screw not by its (infinitesimal) rotation angle, but rather by its (assumed to be finite) rotational velocity.

We now enter into the special circumstances of our top (that is, naturally, of the generalized top). Here we will place the reference point O at the fixed support point. The pitch of the screw is then zero; the screw-motion becomes a simple rotation about an axis passing through O. We thus have the theorems:

An arbitrary motion of our top can be replaced, with respect to its final result, by a rotation with a specific axis, a specific rotation angle, and a specific rotation sense;

and

Every instantaneous (infinitesimal) motion of the top is a rotation with a specific axis, a specific rotational velocity, and a specific sense.

There may next follow some remarks about the composition of rotations, in which we need consider, with respect to the top, only rotations whose axes pass through O. We suppose that the top is given two successive finite rotations. According to the previous theorem, we can effect the result of these two rotations by a single rotation. We obtain the properties of this single rotation from the following theorem:*) if we rotate space successively about the three edges of a three-sided corner, each rotation through twice the corresponding edge angle, then we return to the initial position.[20] This theorem yields the following construction for the single resultant of two given rotations: we place a unit sphere around O, connect its intersection points with the axes of the individual given rotations by a great circle, and apply to this great circle the *half*-angles of the respective individual rotations. *The third corner of the resulting spherical triangle then gives the axis of the resultant rotation, and the adjacent exterior angle gives the half-angle of the resultant rotation.*[21] This construction operates in a remarkable way with the half-angles of rotation, so that the value of the half-angle of the rotation that follows from the construction is determined up to an additive multiple of 2π (that is, the value of the entire angle of the rotation is determined modulo 4π).

We now speak of infinitesimal rotations or rotational velocities. As usual, we assign the infinitesimal rotation a geometric representation

*) Cf. Schell: *Theorie der Bewegung*, Leipzig 1879 II. Teil, Kap. II, §9. The theorem plays a large role in Hamilton's *Lectures on quaternions* (art. 217 and ff.).

by the following procedure: we extend from O, along the axis of rotation, a line segment that represents the magnitude of the rotational velocity, and extends, in particular, in the direction from which the rotation appears to occur in the clockwise sense. We call the resulting geometric counterpart of the infinitesimal rotation a *rotation vector*. If this rotation vector is known, then the axis, velocity, and sense of the infinitesimal rotation follow in an unambiguous way.

We have only to add a convention regarding the unit of measure with which we extend the line, and the system of units with which we wish to measure the angular velocity. It is simplest, here and throughout the following, to adopt the so-called *"absolute system of measure,"* and thus to measure length in centimeters and time in seconds. An angular velocity will always be expressed in arc measure; thus, for example, by the arc of a circle, measured in cm, that a point extending 1 cm from the rotation axis would describe during one second of uniform rotation. In the absolute system of measure, every rotation has, in this sense, a specific numerical value, say n. We determine our representative line segment by extending n cm directly on the rotation axis in the manner given above.

The relevant theorem for the composition of two infinitesimal rotations is now simply:

Two infinitesimal rotations are composed according to the parallelogram law of forces; that is, the corresponding rotation vectors add geometrically (as line segments or vectors).[22]

This fundamental and very well known theorem justifies after the fact the introduction of the word rotation *vector*, and shows, moreover, that the resultant of two infinitesimal rotations is *independent* of their order, and that *infinitesimal rotations are thus interchangeable operations*. For the proof of this it is enough to consider the figure of the parallelogram. We remark that, in contrast, the resultant of two finite rotations changes if we reverse the order of the two rotations, and thus *finite rotations are not interchangeable*. The proof follows from the construction that is indicated on page 10, in which the defining elements of the two rotations are used in an asymmetric manner.[23]

We now consider the moving top in an entire series of different positions, therefore taking a first, second, third, ... of these positions

into view. We replace the motions that lead from the first position to the second, from the second to the third, etc., by single rotations, and thus obtain a series of different rotation axes passing through O. Here we distinguish, as usual in kinematics, a *moving frame* and a *fixed frame*. The moving frame is our top, and the fixed frame is ideal space.

We remark, further, that the distinction between one frame and the other that lies in the words "moving" and "fixed" is actually unjustified from the standpoint of pure kinematics, and that it would be more correct to speak, for example, of a *first* and a *second* frame. Namely, every motion is as equally valid geometrically as its inversion, in which the roles of the moving and fixed frames are interchanged. *Kinematics*, therefore, always treats only of *relative motion*. In *kinetics* it is different. The necessary forces for the generation of a motion change completely if we interchange the moving frame and the fixed frame. In kinetics, therefore, the *direct* and the *inverse motions* have, in general, completely different characters. We will later emphasize an exception to this rule, when we become acquainted with the theorem that the inversion of the motion of the top has, under special circumstances, the same kinetic character as the direct motion.

We wish to mark, in the moving frame and in the fixed frame, the positions of the axes of the above rotations that bring the top from the first position to the second, from the second position to the third, etc. We thus obtain, if we join the successive axes by planes, a *pyramid*[24] fixed in space and a *pyramid* fixed in the top, with equal respective side angles. In the first rotation, the respective first edges of the two pyramids coincide. The moving frame turns around this edge until the second edges become coincident. In the second rotation, the moving frame turns around the second edge. The magnitude of the rotation is such that the respective third edges must coincide at the end of this rotation. So it continues. We can describe the entire rotation process succinctly in the following manner:

The moving pyramid rolls on the fixed pyramid.

This motion must, naturally, necessarily coincide with the actual motion of the top only at each of the final positions of the individual rotations. The intermediate positions can be very different in the two cases. To attain in this manner a complete reproduction of the actual

motion, however, we the make limit passage from finite to infinitesimal rotations. We conceive of the motion of the top at each moment as an infinitesimal rotation, and imagine the corresponding "instantaneous axis of rotation" constructed in the fixed frame as well as in the moving frame. Our pyramids are transformed by the limit passage into cones. We obtain *a cone fixed in space* and *a cone fixed in the top* that have their common vertex at the support point. We denote the cone fixed in the top as the *cone of the polhode* (in German, the path of the rotation pole or the rotation axis), and we call the cone fixed in space the *cone of the herpolhode* (in German, the path along which the rotation pole crawls; since the word is derived from the Greek ἕρπειν, to crawl, one should properly say herpopolhode). During the motion these two cones now roll over one another without slipping, and we have the theorem:

An arbitrary continuous motion of the top can be reproduced completely in all its positions if we let a moving cone roll on a fixed cone. The moving cone is the locus of the instantaneous rotation axis in the top; the fixed cone is the locus of the rotation axis in space.

We have thus reproduced the beautiful and intuitive image with which one represents the motion of the top kinematically according to Poinsot*). This image is very useful whenever the rolling cones have a clear form, particularly in the case of regular precession (cf. the latter sections of this chapter). In more complicated cases, in contrast, where the form of the cones is not directly known, and, if it is, the intuition is of difficult access, the Poinsot representation appears to be somewhat problematic.

Corresponding considerations may also be carried out for the freely moving rigid body. The only difference is that the successive instantaneous axes do not pass through a fixed point, but rather are generally distributed in a twisted manner through space. *As a result, two generalized straight-line surfaces (ruled surfaces) appear here instead of cones.* Also, the instantaneous motion is not a pure rotation, but rather is an infinitesimal screw. *The two ruled surfaces therefore not only roll on each other, but rather slide on each other along their generators in a specific way at the same time; they grind*[25] *on each other, as the engineers say.*

*) Cf. Poinsot: Théorie nouvelle de la rotation des corps, Paris 1834, and Liouv. Journ. sér. 1. t. 16, German translation by Schellbach, Berlin 1851.

We may next go over from the cones of the polhode and the herpolhode to certain curves lying on them. We imagine, for this purpose, the "instantaneous rotation vector" to be placed on the instantaneous rotation axis according to the rule given above. The endpoint of this vector then describes a curve in the fixed frame and a curve in the moving frame, which we call the *herpolhode curve* and the *polhode curve*, respectively.[26] *At the same time that the cones roll during the motion of the top, these two curves roll on each other as well.* Since they express more about the actual rotation than the corresponding cones (namely, they convey the magnitude and the sense of the rotation in addition to the axis), it will frequently be advantageous to make use of these curves in place of the cones.

In the construction of the rotation vectors, we have tacitly assumed that the observer who determines the sense of the rotation and the direction of the rotation vector takes a fixed position in space and observes the motion of the top relative to fixed space. We can also, however, construct the rotation vector of the inverse motion. We must then assume an observer who has a fixed position in the top and observes the rotation of space with respect to the (for him) fixed top. This rotation has, at each instant, the opposite sense of the preceding. The rotation vector of the inverse motion is therefore placed in the direction opposite to that of the direct motion. The endpoint of this vector describes a curve in the fixed frame and a curve in the moving frame, which we can call the *polhode curve and the herpolhode curve of the inverse motion. These curves lie diametrically opposed, with respect to the support point, to the herpolhode curve and the polhode curve, respectively, of the direct motion.*

In conclusion, a place may be found for a remark concerning the experimental verification of the preceding theory. One would wish to observe the existence and position of the instantaneous rotation axis of the moving top directly with the eye. This is difficult, especially for rapid rotation. It can be accomplished, however, by means of an ingenious method given by Maxwell*). For this purpose, Maxwell fixes to the figure axis of the top a pasteboard disk that is divided into four differently colored sectors. During the motion, one sees in the

*) Maxwell, Transact. R. Scott. Soc. of Arts 1855 or Scientific papers I, p. 246.

neighborhood of the rotation axis the color applied there; at some distance, however, the colors of the disk run together. *Thus the instantaneous axis appears as the center point of a colored spot that is surrounded by an undetermined gray. The successive positions of the colored spot make the form of the herpolhode cone directly evident to an observer standing outside the top. The change of the color, moreover, gives an approximate indication of how the instantaneous axis moves with respect to the top.*[27]

§2. Analytic representation of rotations about a fixed point

We must now supplement the geometric considerations of the preceding section through analytic development. For this purpose we use rectangular coordinates, and, in particular, a *coordinate frame xyz fixed in space* and a simultaneous *coordinate frame XYZ moving in space but fixed in the top.*

Both coordinate frames should have their origin at the fixed point O of the top. The unit of measure should be the same for the two frames and for all of the axes. We will usually suppose the z-axis to be extended in the vertical direction, and reckoned as positive upward. As seen from the positive z-axis, the positive x-axis may be transformed into the positive y-axis by a rotation in the clockwise sense. The same stipulation should also apply to the respective positions of the positive X, Y, Z axes. According to this convention, our coordinate frames are designated as *equi-oriented.*[28] For the symmetric top, we will always let the positive Z-axis coincide with the figure axis.

To determine the position of the top in space, it is enough to give the position of the XYZ frame with respect to the xyz frame. This can done if we construct the *transformation formulas* for the passage from one coordinate frame to the other; that is, the formulas that permit the coordinates of an arbitrary point with respect to one frame to be calculated if its coordinates in the other frame are given.

We wish to suppose that the moving frame coincides with the fixed frame in an “initial position”; we imagine every other position to be produced by a rotation from the initial position. We can then interpret the coordinate transformation formulas—as we interpret the transformation equations of analytic geometry in general—in a twofold manner.

1. *We consider a point fixed in space, and ask for its coordinates with*

respect to the moving frame. The point may be given in terms of its coordinates xyz with respect to the fixed coordinate frame. These coordinates are, at the same time, the coordinates of the point with respect to the moving frame in the initial position. *Our formulas determine, in this interpretation, the coordinates* XYZ *of the same space point with respect to the final position of the moving frame.*

2. *We consider a point fixed in the top and moving in space, and ask for its coordinates with respect to the frame fixed in space.* Let the position of the point in the top be given by the coordinates XYZ with respect to the frame fixed in the top. These are, at the same time, the coordinates of the point with respect to the coordinate frame fixed in space in the initial position. *Our formulas yield, in this interpretation, the coordinates* xyz *of the same top point in its final position with respect to the coordinate frame fixed in space.*

In the following, we will have alternative use of both interpretations, which are covered, as is well understood, by one and the same system of formulas. In both cases, we may say, the study of the transformation formulas leads to the *analytic representation of rotations about a fixed point*, where the first interpretation treats of the rotations of a coordinate frame relative to fixed space, and the second interpretation treats of the rotation of space relative to a fixed coordinate frame.

The formulas for the passage from the coordinate frame x, y, z to the coordinate frame X, Y, Z are represented, as is well known, by a so-called *orthogonal* transformation. One has, on the one hand,

$$\begin{aligned} x &= aX + a'Y + a''Z, \\ y &= bX + b'Y + b''Z, \\ z &= cX + c'Y + c''Z, \end{aligned} \tag{1}$$

and, on the other hand,

$$\begin{aligned} X &= a\,x + b\,y + c\,z, \\ Y &= a'x + b'y + c'z, \\ Z &= a''x + b''y + c''z, \end{aligned} \tag{2}$$

where the coefficients in these equations denote the cosines of the angles that the positive half-axes of one frame form with respect to those of the other. Specifically,

$$\begin{aligned} a &= \cos(xX), & a' &= \cos(xY), & a'' &= \cos(xZ), \\ b &= \cos(yX), & b' &= \cos(yY), & b'' &= \cos(yZ), \\ c &= \cos(zX), & c' &= \cos(zY), & c'' &= \cos(zZ). \end{aligned}$$

Equations (2) result from equations (1) through the "transposition of the coefficients." We encompass both sets of equations in the schema

(3)

$$\begin{array}{c|c|c|c} & X & Y & Z \\ \hline x & a & a' & a'' \\ \hline y & b & b' & b'' \\ \hline z & c & c' & c'' \end{array}$$

which when read from left to right reproduces equations (1), and when read from upper to lower reproduces equations (2).

The nine quantities $a, \ldots c''$ are not independent of each other; there is an entire series of relations among them. In fact, as we have said earlier, our top has three degrees of freedom; this means nothing else than that its collected positions can be represented through three independent parameters. We have therefore used *supernumerary parameters* for the representation of rotations in the above transformation equations. The relevant relations among the coefficients are found in the textbooks; we do not stop to give these.[29] We will, in contrast, pose the problem of giving *a representation of rotations through three independent parameters.*

For this purpose one traditionally uses, in the first place, the so-called *asymmetric Euler angles* φ, ψ, ϑ.

In order to be able to define these angles unambiguously, we first explain what we mean (for our symmetric top) by the "*line of nodes.*" We designate as the line of nodes the *half-line* that is simultaneously perpendicular to the vertical and to the figure axis, and as seen from which the vertical is transformed into the figure axis through a clockwise rotation along the shortest path. This specification of the line of nodes will become indeterminate only if the figure axis is in the vertical position; that is, if the axes of z and Z enclose an angle of 0 or π. (In a general orientation of the coordinate systems and for general mass distribution, we have only to say the z- and the Z-axes instead of the vertical and the figure axes in the preceding definition.)

We consider, in addition to our xyz and XYZ frames, an additional third auxiliary coordinate frame that we wish to rotate in three steps from the position xyz to the position XYZ. Specifically, we make the following rotations:

I. Our auxiliary frame is rotated, from its initial position xyz, around the positive z-axis in the clockwise sense, until its positive x-axis

coincides with the line of nodes. We denote the new position of the auxiliary frame by $x_1y_1z_1$; *the angle through which we must turn defines the parameter* ψ.

II. We then rotate the auxiliary frame from its position $x_1y_1z_1$ around the positive x_1-axis (that is, around the line of nodes), again in a clockwise sense, until its positive z-axis coincides with the positive Z-axis. Our auxiliary frame is thus transformed into the position $x_2y_2z_2$. *The angle through which we must rotate in this operation is called* ϑ.

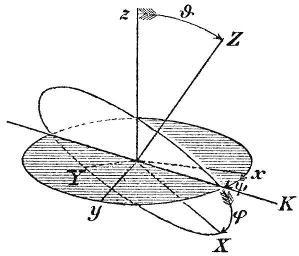


Fig. 3.

III. We rotate, finally, the auxiliary frame about the positive z_2-axis (or, which is the same, about the positive Z-axis), yet again in the clockwise sense, until the positive x_2- and y_2-axes coincide with the positive X- and Y-axes. *The corresponding rotation angle gives us the third parameter* φ. The auxiliary frame has now been transformed to its final position XYZ.

More concisely, but less precisely, we can say (cf. the figure[30]) that the Euler angle

φ	denotes	the	angle	from	the	line of nodes	to	the	pos.	X-axis,
ψ	"	"	"	"	"	"	"	"	"	x-axis,
ϑ	"	"	"	"	"	pos. Z-axis	"	"	"	z-axis.

We will assemble the coefficients of our rotation transformation from these angles φ, ψ, ϑ. For this purpose, we first present the expressions for operations I, II, and III individually.

Operation I leaves the z-axis unchanged, and rotates the xy-plane into itself. We thus have, in the first place, the equation $z_1 = z$. On the other hand, x_1 and y_1 are related to x and y by the transformation equations of a rotation in the plane. We must, to remain in conformity with the conventions concerning the spatial coordinate frames, determine the positive y-axis in such a way that it results from the positive x-axis through a rotation about O in the clockwise sense, which is therefore directly opposite to what is usually done in the analytic geometry of the plane. On the basis of the adjacent figure, the schema for our plane rotation is now

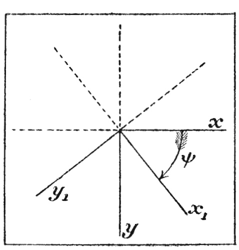


Fig. 4.

(4)

	x_1	y_1
x	$\cos\psi$	$-\sin\psi$
y	$\sin\psi$	$\cos\psi$

The transformation equations for operation I thus become

$$\begin{aligned} x &= \cos\psi\cdot x_1 - \sin\psi\cdot y_1, \\ y &= \sin\psi\cdot x_1 + \cos\psi\cdot y_1, \\ z &= \qquad\qquad\qquad\qquad z_1. \end{aligned}$$

We rewrite these equations by analogy with (3) in the form of a three-by-three schema, and adjoin the correspondingly formed schemata for operations II and III. There follow

I.

	x_1	y_1	z_1
x	$\cos\psi$	$-\sin\psi$	0
y	$\sin\psi$	$\cos\psi$	0
z	0	0	1

II.

	x_2	y_2	z_2
x_1	1	0	0
y_1	0	$\cos\vartheta$	$-\sin\vartheta$
z_1	0	$\sin\vartheta$	$\cos\vartheta$

III.

	X	Y	Z
x_2	$\cos\varphi$	$-\sin\varphi$	0
y_2	$\sin\varphi$	$\cos\varphi$	0
z_2	0	0	1

If we combine these individual transformations, then we obtain the desired transformation equations between the coordinates xyz and XYZ, expressed in terms of the Euler angles in the form

(5)

	X	Y	Z
x	$\cos\varphi\cos\psi-\cos\vartheta\sin\varphi\sin\psi$	$-\sin\varphi\cos\psi-\cos\vartheta\cos\varphi\sin\psi$	$\sin\vartheta\sin\psi$
y	$\cos\varphi\sin\psi+\cos\vartheta\sin\varphi\cos\psi$	$-\sin\varphi\sin\psi+\cos\vartheta\cos\varphi\cos\psi$	$-\sin\vartheta\cos\psi$
z	$\sin\vartheta\sin\varphi$	$\sin\vartheta\cos\varphi$	$\cos\vartheta$

We note, in particular, that the direction cosines of the vertical axis in the moving frame and the direction cosines of the figure axis in the fixed frame have, respectively, the values

$$\left\{\begin{array}{llllll} c &= \sin\vartheta\sin\varphi, & c' &= \sin\vartheta\cos\varphi, & c'' &= \cos\vartheta, \\ a'' &= \sin\vartheta\sin\psi, & b'' &= -\sin\vartheta\cos\psi, & c'' &= \cos\vartheta. \end{array}\right. \tag{6}$$

We now intend to simplify these generally used but, in fact, extremely obscure formulas through an appropriate recapitulation.

This is achieved by first introducing, in place of the *real* coordinates, certain *complex combinations* of them. We set

$$\begin{array}{llll} \xi = & x + iy, & \Xi = & X + iY, \\ \eta = & -x + iy, & \mathsf{H} = & -X + iY, \\ \zeta = & -z, & \mathsf{Z} = & -Z. \end{array}$$

Then the following group of formulas results immediately from the schema (5):

(7)

➳	Ξ	H	Z
ξ	$\dfrac{\cos\vartheta+1}{2}\cdot e^{i(\varphi+\psi)}$	$\dfrac{\cos\vartheta-1}{2}\cdot e^{i(-\varphi+\psi)}$	$i\sin\vartheta\, e^{i\psi}$
η	$\dfrac{\cos\vartheta-1}{2}\cdot e^{i(\varphi-\psi)}$	$\dfrac{\cos\vartheta+1}{2}\cdot e^{i(-\varphi-\psi)}$	$i\sin\vartheta\, e^{-i\psi}$
ζ	$\dfrac{i\sin\vartheta}{2}\cdot e^{i\varphi}$	$\dfrac{i\sin\vartheta}{2}\cdot e^{-i\varphi}$	$\cos\vartheta$

We need not wonder that complex quantities play a role in kinematics. The deeper basis for their appearance will become clear in the following section; it may be pointed out in this place that the equations for a plane rotation that are contained in the schema (4) can also be written to advantage, as is well known, in the complex form*)

$$x \pm iy = e^{\pm i\psi}(x_1 \pm iy_1).$$

The property of the preceding schemata, that they can be read just as well from left to right as from upper to lower, is lost, however, through the introduction of the complex quantities. Our schema (7)

*) It may also be recalled, in this context, that in many areas of mathematical physics, for example in optics, the use of complex quantities (for the symmetric reformulation of otherwise asymmetric equations) is commonplace.

gives only the coefficients in the expressions for $\xi\eta\zeta$ in terms of $\mathsf{\Xi\,H\,Z}$, and may be read only from left to right (as indicated by the attached arrow).

The transformation formulas take an astonishingly simple form, however, if we finally go over from ϑ to $\frac{\vartheta}{2}$, and introduce the abbreviations

$$(8)\qquad \begin{cases} \alpha = \cos\frac{\vartheta}{2}\cdot e^{\frac{i(\varphi+\psi)}{2}}, & \beta = i\sin\frac{\vartheta}{2}\cdot e^{\frac{i(-\varphi+\psi)}{2}}, \\ \gamma = i\sin\frac{\vartheta}{2}\cdot e^{\frac{i(\varphi-\psi)}{2}}, & \delta = \cos\frac{\vartheta}{2}\cdot e^{\frac{i(-\varphi-\psi)}{2}}, \end{cases}$$

which are bound by the relation

$$\alpha\delta - \beta\gamma = +1.$$

Namely, the transformation formulas are then

(9)

$\twoheadrightarrow$	$\mathsf{\Xi}$	H	Z
ξ	α^2	β^2	$2\alpha\beta$
η	γ^2	δ^2	$2\gamma\delta$
ζ	$\alpha\gamma$	$\beta\delta$	$\alpha\delta+\beta\gamma$

The way in which these equations are to be changed if we wish, on the contrary, to express $\mathsf{\Xi\,H\,Z}$ in terms of $\xi\eta\zeta$ (if we therefore wish to replace the horizontal arrow by a vertical arrow) will be discussed on page 31.

The parameters α, β, γ, δ, which will play a large role in the following, are, as one sees, not independent; rather, α and δ, on the one hand, and β and $-\gamma$, on the other hand, are complex conjugates. If we wish to go over to the corresponding four real parameters, then we may set

$$\begin{aligned} \alpha &= D + iC, \qquad & \beta &= -B + iA, \\ \gamma &= B + iA, \qquad & \delta &= D - iC; \end{aligned}$$

we will designate the so-defined parameters A, B, C, D, which are bound by the relation

$$A^2 + B^2 + C^2 + D^2 = 1,$$

as *quaternion quantities*. These quantities provide, in fact, a passage to Hamilton's theory of quaternions. If we now return from $\xi\eta\zeta$ to xyz, the transformation equations in terms of the quaternion quantities

become

(10)

	X	Y	Z
x	$D^2+A^2-B^2-C^2$	$2(AB-CD)$	$2(AC+BD)$
y	$2(AB+CD)$	$D^2-A^2+B^2-C^2$	$2(BC-AD)$
z	$2(AC-BD)$	$2(BC+AD)$	$D^2-A^2-B^2+C^2$

This schema displays a higher degree of symmetry, but is less simple, than the schema (9).

We have thus learnt three different parameter systems for the study of rotations, each of which has its particular merits.

The Euler angles have an immediate geometric meaning; they are advantageous, moreover, for our problem in mechanical respects. (It will later appear that φ and ψ play the roles of "cyclic coordinates" in the theory of the top.) The angles φ, ψ, ϑ corresponding to a rotation are naturally determined only up to an arbitrary additive multiple of 2π. *Our parameters* α, β, γ, δ are the simplest building stones, in analytic respects, from which the formulas for the motion of the top may be formed, as will be evident through the introduction of the elliptic functions. Next to our parameters, the *quaternion quantities* A, B, C, D take a second place; their merit lies in the formal symmetry of calculations. We will return later to the quaternion quantities, as well as to the general position that the theory of quaternions occupies here (cf. especially §7 of this chapter). One notices explicitly that the parameters α, β, γ, δ which correspond to a rotation are determined only up to their signs, as, naturally, are the quantities A, B, C, D. It is enough, in fact, to change φ or ψ by 2π in (8) in order to reverse the signs of α, β, γ, δ. We will later encounter still more particular stipulations in this respect.

The developments of this section extend immediately to the treatment of a freely moving rigid body. In order to determine its position, we can proceed, according to §1, by giving, on the one hand, the position of the reference point O, and, on the other hand, the position of the body relative to O. The former is represented most simply if we prescribe the rectangular coordinates x, y, z of the reference point, the latter if we use one of the three preceding parameter systems.

§3. The meaning of the parameters α, β, γ, δ

We will presume that the simplification of our transformation equations which we have attained through the introduction of the parameters α, β, γ, δ is not accidental, but rather that these quantities have a necessary inner relationship with our problem of rotation. The following remarks should serve to make this relationship understandable, in so far as this is possible without assuming further prior knowledge. It is convenient here to adopt the second of the two interpretations characterized on page 16, and thus to consider a point moving in space with respect to a coordinate system fixed in space.

In a rotation about O, each point of space XYZ is transformed to another point that has the same distance from O. In particular, the points that have distance zero from O must be transformed into the same such points. One can object here that there is no point with this property other than O itself. This is correct, as long as one remains among real points. However, it has long been customary in geometry to consider points with imaginary coordinates, through which the simplicity and clarity of the theory are very significantly enhanced. If we do this, then we must say: points that have distance zero from O are at hand in infinite quantity; they lie on an (imaginary) cone of the second order, whose equation is

$$x^2 + y^2 + z^2 = 0,$$

the so-called *minimal cone.*

We write this equation, for our purpose, in the form

$$(x + iy)(-x + iy) = z^2,$$

or, with the use of the quantities introduced on page 20,

$$\xi\eta = \zeta^2.$$

We then define two parameters λ_1 and λ_2 through the equations

$$\xi = \lambda_1^2, \quad \eta = \lambda_2^2, \quad \zeta = \lambda_1\lambda_2. \tag{1}$$

Evidently, there corresponds to each pair of values λ_1, λ_2 a completely determined point of our imaginary cone (for which the equation $\xi\eta = \zeta^2$ is identically fulfilled by virtue of (1)); in contrast, there correspond to a point on the cone two pairs of values λ_1, λ_2 that differ, however, only by a common change of sign. *The equations* (1), we can say, *give us a parametric representation of the points of our imaginary cone.*

We obtain a second parametric representation for the same set of points if, starting from the equation

$$X^2 + Y^2 + Z^2 = 0,$$

we introduce two parameters $\mathsf{\Lambda}_1$, $\mathsf{\Lambda}_2$ through the equations

$$\mathsf{\Xi} = \mathsf{\Lambda}_1^2, \quad \mathsf{H} = \mathsf{\Lambda}_2^2, \quad \mathsf{Z} = \mathsf{\Lambda}_1\mathsf{\Lambda}_2, \tag{2}$$

where

$$\mathsf{\Xi} = X + iY, \quad \mathsf{H} = -X + iY, \quad \mathsf{Z} = -Z.$$

Since our rotation transforms the points (2) into the points (1), the parameters $\mathsf{\Lambda}_1$, $\mathsf{\Lambda}_2$ will have, at the same time, a certain relation to the parameters λ_1, λ_2. We read the relation between the two pairs of values from the schema (9) of the previous section. We have

$$\begin{aligned}
\lambda_1^2 \quad &= \alpha^2\mathsf{\Lambda}_1^2 + \beta^2\mathsf{\Lambda}_2^2 + 2\alpha\beta\mathsf{\Lambda}_1\mathsf{\Lambda}_2 &&= (\alpha\mathsf{\Lambda}_1 + \beta\mathsf{\Lambda}_2)^2,\\
\lambda_2^2 \quad &= \gamma^2\mathsf{\Lambda}_1^2 + \delta^2\mathsf{\Lambda}_2^2 + 2\gamma\delta\mathsf{\Lambda}_1\mathsf{\Lambda}_2 &&= (\gamma\mathsf{\Lambda}_1 + \delta\mathsf{\Lambda}_2)^2,\\
\lambda_1\lambda_2 &= \alpha\gamma\mathsf{\Lambda}_1^2 + \beta\delta\mathsf{\Lambda}_2^2 + (\alpha\delta + \beta\gamma)\mathsf{\Lambda}_1\mathsf{\Lambda}_2 &&= (\alpha\mathsf{\Lambda}_1 + \beta\mathsf{\Lambda}_2)(\gamma\mathsf{\Lambda}_1 + \delta\mathsf{\Lambda}_2).
\end{aligned}$$

But from this there follows

$$\begin{cases} \lambda_1 = \alpha\mathsf{\Lambda}_1 + \beta\mathsf{\Lambda}_2,\\ \lambda_2 = \gamma\mathsf{\Lambda}_1 + \delta\mathsf{\Lambda}_2, \end{cases} \tag{3}$$

where it is still permitted to reverse the signs of α, β, γ, δ simultaneously.

We see, therefore, that with each orthogonal transformation of the rectangular coordinates xyz, a second homogeneous linear transformation runs in parallel for the parameters defined on our imaginary cone.[31] *The coefficients of this second transformation are precisely our quantities* $\alpha\beta\gamma\delta$. Since $\alpha\delta - \beta\gamma = 1$ was assumed earlier, *this second transformation should have determinant* 1. Here lies a new and simple definition of the parameters α, β, γ, δ. It is in agreement with the concluding remark of the preceding section that the signs of α, β, γ, δ remain undetermined according to this definition. But it is of value, moreover, to cast the new definition of α, β, γ, δ into yet another different form.

We first wish to go over from the points of the cone to the lines running on the cone, its "generators." While we cover the twofold multiplicity of points by two parameters λ_1, λ_2, we can distinguish the onefold multiplicity of generators through a single parameter. Indeed, the parameter

$$\lambda = \frac{\lambda_1}{\lambda_2} = \frac{\xi}{\zeta} = \frac{\zeta}{\eta} \tag{4}$$

offers itself for this purpose. In fact, these equations determine, for each value of λ, an (imaginary) line on the cone, and vice versa. If we define, correspondingly, the parameter

$$\mathsf{\Lambda} = \frac{\mathsf{\Lambda}_1}{\mathsf{\Lambda}_2} = \frac{\Xi}{\mathsf{Z}} = \frac{\mathsf{Z}}{\mathsf{H}},$$

then we have, according to (3), the relation

$$\lambda = \frac{\alpha\mathsf{\Lambda} + \beta}{\gamma\mathsf{\Lambda} + \delta} \tag{5}$$

between λ and $\mathsf{\Lambda}$.

In a rotation about a fixed point, the parameter through which we characterized the generators of our imaginary cone therefore undergoes a fractional linear transformation whose coefficients, in so far as we set the determinant of the transformation equal to 1*, are our quantities* α, β, γ, δ.[32]

But we will also show that we can likewise assign a parameter to all real lines (or, more correctly, half-lines) extending from O, and that this parameter is transformed by the rotation in the same way.

We begin, for this purpose, from the trivial equation

$$x^2 + y^2 + z^2 = r^2, \quad \text{or} \quad \sqrt{x^2 + y^2 + z^2} = r,$$

which is nothing other than the definition of r. With respect to the sign of the root, we agree that this sign should always be calculated as *positive* for real values of x, y, z. We give the above equation, in the same way as earlier, the form

$$(x + iy)(-x + iy) = (z + r)(z - r),$$

or

$$\xi\eta = (\zeta - r)(\zeta + r),$$

or, finally,

$$\frac{\xi}{\zeta + r} = \frac{\zeta - r}{\eta}.$$

If we denote the common value of the right and left sides in the last equation by λ, then we can write

$$\frac{\xi}{\zeta + r} = \lambda, \quad \frac{\eta}{\zeta - r} = \frac{1}{\lambda}. \tag{6}$$

In this manner we have associated to all points of space a (generally complex) parameter λ that is uniquely determined for real points of space as a consequence of our agreement concerning the sign of r. A given parameter value λ is thus found at all points of one and only one real half-line through O. In fact, the value of λ remains unchanged if we

multiply the coordinates x, y, z of a real point of space by a common *positive* factor. The value of λ changes, in contrast, not only if we change the ratios of these quantities, but also if we multiply by a common negative factor, the latter in consequence of our agreement concerning the sign of r. The quantity λ is then transformed to a value λ', which is determined by the equations

$$(6') \qquad \frac{\xi}{\zeta - r} = \lambda', \quad \frac{\eta}{\zeta + r} = \frac{1}{\lambda'};$$

this value is evidently conjugate to the value $-\dfrac{1}{\lambda}$.

Through our equations (6) *and* (6′)*, each real half-line of space is assigned one parameter value, and each real full line is assigned two parameter values.* We remark here that our present parameter representation, extended to the imaginary rays of the cone $r^2 = 0$, is transformed into the parameter representation earlier defined by equation (4), and that, in this case, the two values λ and λ' become identical.

In a corresponding manner, two parameters $\mathsf{\Lambda}$ and $\mathsf{\Lambda}'$ may be defined that cover the real half and full lines in the initial position. Bearing in mind that the distance r remains unchanged by the rotation, so that $R = r$, we set

$$(7) \qquad \frac{\mathsf{\Xi}}{\mathsf{Z} + r} = \mathsf{\Lambda}, \quad \frac{\mathsf{\Xi}}{\mathsf{Z} - r} = \mathsf{\Lambda}', \quad \text{etc.}$$

To now establish the dependence of the parameters λ, λ' and $\mathsf{\Lambda}$, $\mathsf{\Lambda}'$, we express the coordinates ξ, η, ζ and $\mathsf{\Xi}$, H, Z in terms of these parameters. Equations (6) and (6′) give

$$\lambda\lambda' = \frac{\xi^2}{\zeta^2 - r^2} = \frac{\xi^2}{\xi\eta} = \frac{\xi}{\eta}$$

and

$$\lambda + \lambda' = \frac{2\xi\zeta}{\zeta^2 - r^2} = \frac{2\xi\zeta}{\xi\eta} = \frac{2\zeta}{\eta}.$$

We can thus set

$$\xi : \zeta : \eta = \lambda\lambda' : \frac{\lambda + \lambda'}{2} : 1,$$

or, with the use of a proportionality factor ϱ,

$$(8) \qquad \xi = \varrho\lambda\lambda', \quad \zeta = \varrho\,\frac{\lambda + \lambda'}{2}, \quad \eta = \varrho.$$

In the same way, there follow from equations (7)

$$(9) \qquad \mathsf{\Xi} = \sigma\mathsf{\Lambda}\mathsf{\Lambda}', \quad \mathsf{Z} = \sigma\,\frac{\mathsf{\Lambda} + \mathsf{\Lambda}'}{2}, \quad \mathsf{H} = \sigma,$$

where σ denotes a second proportionality factor.

With the abbreviation $\frac{\varrho}{\sigma} = \tau$, our transformation formulas (9) on page 21 take the form

$$\begin{aligned}
\tau \cdot \lambda\lambda' &= \alpha^2 \Lambda\Lambda' + \alpha\beta(\Lambda + \Lambda') + \beta^2 = (\alpha\Lambda + \beta)(\alpha\Lambda' + \beta), \\
\tau \qquad &= \gamma^2 \Lambda\Lambda' + \gamma\delta\,(\Lambda + \Lambda') + \delta^2 = (\gamma\Lambda + \delta)(\gamma\Lambda' + \delta), \\
\tau \cdot \frac{\lambda + \lambda'}{2} &= \alpha\gamma\Lambda\Lambda' + (\alpha\delta + \beta\gamma)\,\frac{\Lambda + \Lambda'}{2} + \beta\delta \\
&= \frac{1}{2}\,(\alpha\Lambda + \beta)(\gamma\Lambda' + \delta) + \frac{1}{2}\,(\gamma\Lambda + \delta)(\alpha\Lambda' + \beta).
\end{aligned}$$

Through appropriate division there follow

$$\begin{aligned}
\lambda\lambda' &= \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta} \cdot \frac{\alpha\Lambda' + \beta}{\gamma\Lambda' + \delta}, \\
\lambda + \lambda' &= \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta} + \frac{\alpha\Lambda' + \beta}{\gamma\Lambda' + \delta}.
\end{aligned}$$

If we imagine the parameter value Λ of our half-line to be given before the rotation, then its parameter λ after the rotation is determined, in the first place, only in a double-valued way. We have either

$$\lambda = \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta}, \qquad \lambda' = \frac{\alpha\Lambda' + \beta}{\gamma\Lambda' + \delta},$$

or

$$\lambda = \frac{\alpha\Lambda' + \beta}{\gamma\Lambda' + \delta}, \qquad \lambda' = \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta}.$$

One notes that both formulas coincide with (5) if one sets $\lambda = \lambda'$, $\Lambda = \Lambda'$, as holds for the points of the minimal cone.

But now it is clear that only one of these two relations can signify a rotation. Namely, the second row is obtained from the first by the interchange of λ with λ', and thus by transforming each half line into its opposite. This operation, however, is not possible by any motion of three-dimensional space. On the other hand, one notices that each of the two formulas represents a *continuum* of transformations. We can therefore say immediately that the operation represented by one row transforms the bundle of half-lines in like orientation, and the operation represented by the other row transforms the half-lines in opposite orientation.

We can recognize most easily which of the two rows corresponds to a rotation if we go over to a special case. The simplest case of a rotation is the "null rotation," in which each point of space is transformed into itself. The null rotation corresponds to the parameter values $\vartheta = \varphi = \psi = 0$, and therefore to $\alpha = \delta = \pm 1$ and $\beta = \gamma = 0$. The first row

yields, in this case, $\lambda = \Lambda$, as must be true; the second row, however, gives $\lambda = \Lambda'$, which is impossible. Therefore only the first row represents a rotation; we necessarily have

$$\lambda = \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta}, \tag{10}$$

and we state the theorem:

If we assign each real half-line through O a parameter λ by equation (6), *then this parameter is transformed under a rotation about O by a fractional linear transformation, exactly like the parameter λ through which we earlier distinguished the generators of our imaginary cone.* We can also, therefore, take this as the definition of α, β, γ, δ. —

If we had wished to appeal to prior knowledge of projective geometry for these developments, then we could have simplified our considerations essentially. We would first have individualized, as above, the generators of the minimal cone through a parameter λ. Since the minimal cone is transformed into itself under a rotation, this parameter undergoes a projective transformation. This yields the theorem of formula (5), from which one ascends directly to formula (3). We would then have referred each full line through O to the parameters λ of those two generators that rest in the tangent plane to the minimal cone passing through the line. We would thus come correctly to the definitions of the values of λ and λ' that are contained in equations (6).[33] That we assign λ to one half-line and λ' to the other is an incidental convention that we establish on no particular basis. The latter theorem, which states that these values transform projectively under a rotation, is now obviously an immediate consequence of (5), since λ and λ' are now parameters of the minimal cone. —

Finally, we bring our parameter representation into an interesting relation with a manner of presentation that has been common in function theory since the time of R i e m a n n. Namely, we wish to collect the system of our half-lines by means of a spherical surface of radius 1 that is described around O. Each point of the sphere will thus be assigned a specified half-line and therefore a specified value of the parameter λ. *The complex quantity λ is therefore distributed uniquely on the spherical surface*, and the point xyz with $x^2 + y^2 + z^2 = 1$ corresponds, according to (6), to the quantity

$$\lambda = \frac{x + iy}{1 - z}. \tag{11}$$

We denote the intersection points of the sphere with the positive and negative z-axis as the north and south poles, respectively. We then have the value $\lambda = \infty$ at the north pole ($z = +1$) and the value $\lambda = 0$ at the south pole ($z = -1$); the real values of λ lie on the meridian $y = 0$, and the equator $z = 0$ contains those points for which $|\lambda| = 1$.

In function theory, one calls this parameter assignment the *interpretation of the complex variable λ on the Riemannian spherical surface.*

One usually arrives at this interpretation by first interpreting the parameter λ in the plane according to G a u ſ s, and specifically in the equatorial plane of our unit sphere, and then mapping this plane through stereographic projection from the north pole to the sphere. For us it would be highly advisable to take the opposite path. After we have arrived directly at the definition of the complex variable on the *Riemannian spherical surface* through the preceding considerations, which arose spontaneously from the nature of our problem, it can, after the fact, be desirable for certain purposes to go over to the *Gauſsian plane.* If, for example, we have discussed a curve on the spherical surface with the help of the variable λ and now wish to represent this curve graphically, then we must somehow refer the sphere to a displaying plane. The stereographic projection serves as the best procedure for this purpose, and with it comes the passage to the Gauſsian plane. This is simple analytically, in that we set $\lambda = u + iv$ and interpret u and v as rectangular coordinates in the plane.

Just as we carry over the parameter λ from the system of half-lines to the Riemannian spherical surface, so the meaning of α, β, γ, δ naturally carries over naturally from the half-lines to this surface.

We consider two commonly placed unit spheres about O, of which we imagine one as fixed in space, and the other as moving. We extend in the prescribed way the parameter λ on the fixed sphere and the parameter Λ on the moving sphere, where, in the initial position, each two coinciding points of the spheres carry coinciding values of λ and Λ. *If we now apply an arbitrary rotation to the moving sphere, then the parameters λ and Λ, which correspond to these two coinciding points, have, after the rotation, the relation*

$$\lambda = \frac{\alpha\Lambda + \beta}{\gamma\Lambda + \delta}, \tag{12}$$

where α, β, γ, δ are precisely the earlier defined quantities. If, in contrast, we let formula (12) be the definition of α, β, γ, δ, we must only add that the determinant $(\alpha\delta - \beta\gamma)$ should always equal 1. The

definition derived in this way has, in preference to the earlier (3), (5), and (10), the advantage of greater clarity, and thus we will generally take it as a basis in the following.

The idea that rotations in space can be represented by fractional linear transformations of the complex variable λ, which appears incidentally in R i e m a n n*), comes to full value in the modern theory of regular bodies**).

§4. The use of α, β, γ, δ for the study of finite rotations

When we study, in the following, the properties of the rotation transformation, we will usually carry out the calculations in our parameters α, β, γ, δ, where they become simplest. Instead of operating with the transformation equations for x, y, z, we generally prefer to make use of the device of operating with the simpler transformation equations for the quantities λ_1, λ_2, which, as we saw in the preceding section (cf. eqn. 3), are necessarily linked to those for x, y, z.

In this spirit, we first calculate the *inverse* of a given rotation. If the given rotation transforms a point fixed in the top from the initial position XYZ to the final position xyz, then the inverse rotation is the rotation that transforms the point from the position xyz back to XYZ. The given rotation has the parameters α, β, γ, δ; it is asked what the parameters of the inverse rotation are.

We consider

$$\lambda_1 = \alpha\Lambda_1 + \beta\Lambda_2,$$
$$\lambda_2 = \gamma\Lambda_1 + \delta\Lambda_2;$$

the inverse transformation follows from these equations by algebraic solution. There follows, with consideration of $\alpha\delta - \beta\gamma = 1$,

$$(1) \qquad \begin{cases} \Lambda_1 = \delta\lambda_1 - \beta\lambda_2, \\ \Lambda_2 = -\gamma\lambda_1 + \alpha\lambda_2. \end{cases}$$

We thus obtain the inverse transformation from the given transformation if we interchange α and δ and reverse the signs of β and γ.

We carry over this result to our other parameter systems A, B, C, D and φ, ψ, ϑ. First, one recognizes immediately from the definition of the quaternion quantities on page 21:

If a rotation is given to us by the values of A, B, C, D, then we obtain the inverse rotation if we reverse the signs of A, B, and C but leave D unchanged.

---

*) Cf. the treatise: Über die Flächen von kleinstem Inhalte bei gegebener Begrenzung, art. 8. Ges. W., 2. Aufl. p. 309.

**) Klein: Vorlesungen über das Ikosaeder, Cap. II, §1.

Further, there follows from the relation of α, β, γ, δ to φ, ψ, ϑ (cf. page 21), or also directly from the geometric meaning of the latter (cf. page 17 f.):

If the direct rotation is given by the angles φ, ψ, ϑ, then the inverse rotation is determined by the angles $-\psi$, $-\varphi$, $-\vartheta$.

We apply these results to solve the equations contained in the schema (9) of page 21 for $\mathsf{\Xi}$, H, Z. For this purpose, we need only enter δ, $-\beta$, $-\gamma$, α there instead of α, β, γ, δ, respectively. The solution of the schema (9) is thus

$\twoheadrightarrow$	ξ	η	ζ
$\mathsf{\Xi}$	δ^2	β^2	$-2\beta\delta$
H	γ^2	α^2	$-2\alpha\gamma$
Z	$-\gamma\delta$	$-\alpha\beta$	$\alpha\delta+\beta\gamma$

(where the attached arrow signifies that this schema should be read only from left to right). If we insert, on the other hand, $-A$, $-B$, $-C$, $+D$ or $-\psi$, $-\varphi$, $-\vartheta$ instead of A, B, C, D or φ, ψ, ϑ, respectively, in the schema (10) or (5) of the section before the last, then we obtain a previously known result: there follows only a transposition of the coefficients, as is obvious, according to page 17, for the inverse of an orthogonal transformation.

In a corresponding manner, we treat of the *composition of two rotations*. We consider a rotation with parameters α, β, γ, δ and a second rotation with parameters α', β', γ', δ'. It is then desired to compute the parameters α'', β'', γ'', δ'' that, in their application, are equivalent to the composition of the first and second rotations. It will be convenient here to adopt the first of the two interpretations of page 16, and therefore to consider the coordinates of a space point with respect to a successively rotated coordinate system. In particular, we have three positions of the moving coordinate system to distinguish: 1) an initial position, in which the moving system coincides with the fixed system, 2) a position into which the moving system is transformed through the first rotation, and 3) a final position, into which the moving system is transformed from position 2) by the second rotation. Let the space point be given by its coordinates xyz in the fixed coordinate system. Then xyz are at the same time the coordinates of the space point before the first rotation with respect to position 1) of the moving coordinate system. Further, let $x'z'y'$ and XYZ be the coordinates of the same

space point with respect to positions 2) and 3), respectively, of the moving coordinate system.

We write, in an easily understood symbolism,

$$(x\ y\ z\ ) = (\alpha\ \beta\ \gamma\ \delta\ )\ (x'y'z'),$$
$$(x'\,y'\,z'\,) = (\alpha'\,\beta'\,\gamma'\,\delta'\,)\ (XYZ).$$

We seek the parameters α'', β'', δ'', γ'' in the symbolic equation

$$(xyz) = (\alpha''\beta''\delta''\gamma'')(XYZ).$$

If we go over to the complex quantities λ, λ', Λ that correspond to the coordinates xyz, $x'y'z'$, XYZ, respectively, then our equations mean the following:

$$\left\{\begin{aligned} \lambda_1 &= \alpha\lambda_1' + \beta\lambda_2', \\ \lambda_2 &= \gamma\lambda_1' + \delta\lambda_2', \end{aligned}\right. \tag{2}$$

$$\left\{\begin{aligned} \lambda_1' &= \alpha'\Lambda_1 + \beta'\Lambda_2, \\ \lambda_2' &= \gamma'\Lambda_1 + \delta'\Lambda_2, \end{aligned}\right. \tag{2'}$$

and

$$\left\{\begin{aligned} \lambda_1 &= \alpha''\Lambda_1 + \beta''\Lambda_2, \\ \lambda_2 &= \gamma''\Lambda_1 + \delta''\Lambda_2. \end{aligned}\right. \tag{3}$$

However, there follow from equations (2) and (2′), by elimination,

$$\lambda_1 = (\alpha\alpha' + \beta\gamma')\Lambda_1 + (\alpha\beta' + \beta\delta')\Lambda_2,$$
$$\lambda_2 = (\gamma\alpha' + \delta\gamma')\Lambda_1 + (\gamma\beta' + \delta\delta')\Lambda_2.$$

The comparison of these equations with equations (3) yields the desired values of the parameters α'', β'', γ'', δ''; namely,

$$\left\{\begin{aligned} \alpha'' &= \alpha\alpha' + \beta\gamma', \qquad \beta'' = \alpha\beta' + \beta\delta', \\ \gamma'' &= \gamma\alpha' + \delta\gamma', \qquad \delta'' = \gamma\beta' + \delta\delta'. \end{aligned}\right. \tag{4}$$

We verify that these quantities are *bilinear two-term* combinations of $\alpha\beta\gamma\delta$ and $\alpha'\beta'\gamma'\delta'$, and that the latter quantities enter our formulas *asymmetrically.* We thus state the theorem:

If we successively apply two rotations and determine the rotation that is equivalent to these two together, then the parameters α, β, γ, δ of the new rotation will be bilinear two-term combinations of the parameters of the given rotations. We must give the order of the combining rotations with care: if we reverse their order, the result of their composition changes.

We will wish to carry over the preceding composition rule to the other parameter systems as well. With respect to the quaternion quan-

tities A, B, C, D, this is done by simply separating equations (4) into their real and imaginary parts. If A, A', A'' etc. correspond to the quantities α, α', α'' etc., respectively, then there follow

$$\text{(5)}\qquad \begin{cases} A'' = AD' + BC' - CB' + DA', \\ B'' = -AC' + BD' + CA' + DB', \\ C'' = AB' - BA' + CD' + DC', \\ D'' = DD' - AA' - BB' - CC'. \end{cases}$$

The parameters of the resulting rotation are again bilinear in the parameters of the individual rotations. However, the expressions have now become four-term.

We will not write the corresponding formulas for φ, ψ, ϑ explicitly, since they are rather complicated. They result from equations (4) if we express α, β, γ, δ in terms of φ, ψ, ϑ, as given on page 21.

We now calculate the *characteristic geometric elements of a rotation; namely, the position of the rotation axis and the magnitude of the rotation angle*, assuming that the rotation parameters α, β, γ, δ are known. This does not come without some calculation. To obtain an appropriate starting point, we first imagine the rotation angle and rotation axis as known. We make the rotation axis the first coordinate axis of a rectangular coordinate system that is fixed in space with origin O, in which U, V, W are the coordinates of a point before the rotation, and u, v, w are the coordinates of the same point after the rotation. In this coordinate system the formulas for the rotation transformation become particularly simple. They are contained in the schema II of page 19; if we denote the rotation angle by ω and use, as on page 20, the complex combinations $V + iW$, $V - iW$ instead of the coordinates V, W, etc., then the transformation formulas become

$$\text{(6)}\qquad \begin{cases} u = U, \\ v + iw = e^{i\omega}(V + iW), \\ v - iw = e^{-i\omega}(V - iW). \end{cases}$$

We thus conclude the following: the rotation axis, or, as we prefer to say, the "*primary axis of the rotation*," is distinguished by the fact that its points remain unchanged throughout by the rotation. In addition to this "primary axis," however, there are, if we once again consider imaginary points and lines, two other noteworthy "*secondary axes*," namely

the rays $V+iW=0$ and $V-iW=0$ lying in the plane $U=0$. They are distinguished by the fact that the complex coordinates corresponding to their points change only by a factor. This factor is equal to $e^{-i\omega}$ along the first secondary axis, equal to $e^{+i\omega}$ along the second secondary axis, and, as we may add, equal to 1 along the primary axis. In fact, if we set $U=0$ and $V+iW=0$, then, according to (6), $u=0$ and $v+iw=0$, while $v-iw = e^{-i\omega}(V-iW)$. All three coordinates U, $V+iW$, $V-iW$, we can say, are multiplied along this first secondary axis by the common factor $e^{-i\omega}$. In a corresponding manner, the coordinates U, $V+iW$, $V-iW$ along the second secondary axis are multiplied by the common factor $e^{+i\omega}$, and the coordinates on the primary axis are multiplied by 1.

The same also holds, naturally, for the coordinates U, V, W themselves, and furthermore for all homogeneous linear functions of these quantities; thus, for example, for the coordinates X, Y, Z in an arbitrary rectangular coordinate system that has a common origin with the U, V, W system, and finally also for our previously used quantities $\mathsf{\Xi}$, H, Z. *The coordinates* $\mathsf{\Xi}$, H, Z *that correspond to a point on one of our three rotation axes are also multiplied in the rotation by a common factor; namely, by* 1, $e^{+i\omega}$, and $e^{-i\omega}$, *respectively.* If we denote any of these factors by m, then we have, for the points of our three axes,

$$\xi = m\mathsf{\Xi}, \quad \eta = m\mathsf{H}, \quad \zeta = m\mathsf{Z}. \tag{7}$$

To establish the connection with α, β, γ, δ, we draw upon the schema (9) of page 21. For a point of our rotation axes, the schema (9) implies, with consideration of (7), the equations

$$\begin{cases} 0 = (\alpha^2 - m)\mathsf{\Xi} + \beta^2\mathsf{H} + 2\alpha\beta\mathsf{Z}, \\ 0 = \gamma^2\mathsf{\Xi} + (\delta^2 - m)\mathsf{H} + 2\gamma\delta\mathsf{Z}, \\ 0 = \alpha\gamma\mathsf{\Xi} + \beta\delta\mathsf{H} + (\alpha\delta + \beta\gamma - m)\mathsf{Z}. \end{cases} \tag{8}$$

If these three equations are to be compatible with each other, then their determinant must vanish. The factor m must therefore satisfy the cubic equation

$$\begin{vmatrix} \alpha^2 - m & \beta^2 & 2\alpha\beta \\ \gamma^2 & \delta^2 - m & 2\gamma\delta \\ \alpha\gamma & \beta\delta & \alpha\delta + \beta\gamma - m \end{vmatrix} = 0.$$

In the calculation of the determinant, the coefficients simplify significantly by virtue of the relation $\alpha\delta - \beta\gamma = 1$; our equation takes the form

$$m^3 - (m^2 - m)((\alpha+\delta)^2 - 1) - 1 = 0.$$

The roots of this equation must be identical with the quantities 1, $e^{i\omega}$, $e^{-i\omega}$. In fact, $(m-1)$ may first be removed as a factor. The remaining quadratic equation is

$$m^2 + m(2 - (\alpha+\delta)^2) + 1 = 0.$$

Its roots are

$$m = \frac{(\alpha+\delta)^2}{2} - 1 \pm \sqrt{-(\alpha+\delta)^2 + \frac{(\alpha+\delta)^4}{4}} = e^{\pm i\omega}.$$

Thus there follows

$$\cos\omega = \frac{(\alpha+\delta)^2}{2} - 1,$$

from which the magnitude of the rotation angle is found. The formula simplifies through the introduction of the half-angle; we have

$$\pm\cos\frac{\omega}{2} = \frac{\alpha+\delta}{2} \tag{9}$$

and

$$\pm\sin\frac{\omega}{2} = \sqrt{1 - \left(\frac{\alpha+\delta}{2}\right)^2}. \tag{9'}$$

We will presently go into more detail concerning the determination of the signs in these formulas.

The *magnitude of the rotation angle* having been found, there remains to calculate the *position of the rotation axis* (or, in our above designation, the "primary rotation axis"). Here equations (8) serve us. If we place $m = 1$ in them, then one of the equations will be dispensable because of the vanishing determinant. We choose for the determination of the rotation axis, for example, the two equations

$$\begin{aligned} (\alpha^2 - 1)\Xi + \beta^2\mathsf{H} + 2\alpha\beta\mathsf{Z} &= 0, \\ \gamma^2\Xi + (\delta^2 - 1)\mathsf{H} + 2\gamma\delta\mathsf{Z} &= 0, \end{aligned}$$

from which there follows, by a small determinant calculation,

$$\Xi : \mathsf{H} : \mathsf{Z} = -2\beta : 2\gamma : (\alpha - \delta). \tag{10}$$

We wish to calculate straightaway the direction cosines $\cos a$, $\cos b$, $\cos c$, which our rotation axis encloses with the coordinate axes. Since we assume that the moving system coincides with the fixed system before the rotation, the direction cosines in one system will obviously be identical with those of the other. And, in particular, we have, along the rotation axis,

$$\Xi : \mathsf{H} : \mathsf{Z} = \xi : \eta : \zeta = \cos a + i\cos b : -\cos a + i\cos b : -\cos c.$$

Denoting by ϱ a factor of proportionality, we set, with regard to (10),

$$
(11) \qquad \begin{cases} \cos a + i\cos b = & -2\beta\varrho, \\ -\cos a + i\cos b = & 2\gamma\varrho, \\ -\cos c = & (\alpha-\delta)\varrho. \end{cases}
$$

From the equation $\cos^2 a + \cos^2 b + \cos^2 c = 1$, there follows for ϱ the further condition

$$
\varrho^2\Big((\alpha-\delta)^2 + 4\beta\gamma\Big) = \varrho^2\Big((\alpha+\delta)^2 - 4\Big) = 1,
$$

or, with consideration of $(9')$,

$$
(12) \qquad \varrho = \frac{\pm i}{2\sin\frac{\omega}{2}}.
$$

We now wish to resolve the indeterminacy of the signs in the above and the previous formulas, in so far as this is possible.

First, it is clear that if we consider only the initial and final positions of the body, each angle $\omega + 2k\pi$, where k is an arbitrary integer, can be regarded as the rotation angle instead of the angle ω. In fact, the addition of one or more full rotations does not change the final position of the body at all. From this point of view, therefore, the signs of $\cos\frac{\omega}{2}$ and $\sin\frac{\omega}{2}$ necessarily remain undetermined.

We now wish to stipulate, to the contrary, that we will evaluate the magnitude of a rotation not merely from the initial and final positions of the body, and that we therefore imagine ω to be given not only modulo 2π. *We wish, rather, to regard the intermediate positions of a rotation as known to the extent that we can also fix* $\frac{\omega}{2}$ *modulo* 2π, *and therefore* ω *modulo* 4π.[34]

In spite of this stipulation, however, there still remains an indeterminacy which we cannot well remove. Namely, we can replace, considering the final as well as the intermediate positions of the body, each rotation ω about a specified half-line in the positive sense by a rotation about the opposite half-line in the negative sense. This circumstance corresponds to the fact that we can simultaneously exchange a, b, c, and ω with $a+\pi$, $b+\pi$, $c+\pi$, and $-\omega$.

From what has been said, it follows, on the basis of our stipulation, that $\cos\frac{\omega}{2}$ *and the products* $\sin\frac{\omega}{2}\cos a$ *etc. will receive a determined sign, but that the signs of* $\sin\frac{\omega}{2}$, $\cos a$, $\cos b$, $\cos c$ *taken individually still remain undetermined.*

After this preparation, we return to equations (9) through (12); we now wish to adopt a special position of the coordinate system. We take the rotation axis as the z-axis, and consider a rotation in the clockwise sense, as seen from the positive z-axis, whose magnitude is given through a specified value of $\frac{\omega}{2}$ modulo 2π. We then have $\cos a = \cos b = 0$, $\cos c = 1$; on the other hand, we can give the parameters α, β, γ, δ for this rotation uniquely. Namely, it follows from the definition of the Euler angles that $\vartheta = 0$, $\varphi + \psi = \omega$, or, more precisely said, $\frac{\vartheta}{2} \equiv 0$, $\frac{\varphi+\psi}{2} \equiv \frac{\omega}{2}$ modulo 2π. Thus α, β, γ δ receive the values $\alpha = e^{\frac{i\omega}{2}}$, $\beta = \gamma = 0$, $\delta = e^{-\frac{i\omega}{2}}$. The third of equations (11) now gives

$$-1 = \left(e^{\frac{i\omega}{2}} - e^{-\frac{i\omega}{2}}\right)\varrho\,;$$

that is,

$$\varrho = \frac{+i}{2\sin\frac{\omega}{2}}.$$

At the same time,

$$\frac{\alpha+\delta}{2} = +\cos\frac{\omega}{2}.$$

We will also, correspondingly, choose the two *upper* signs in equations (9) and (12) for a general position of the coordinate system. The signs in equation (9′) remain, in contrast, still undetermined, which, according to the above discussion, is not to be otherwise expected. In order to bring together the developed formulas clearly, we write

$$(13)\quad \begin{cases} \cos a + i\cos b = -\dfrac{\beta i}{\sin\frac{\omega}{2}}, & \\ & \cos\dfrac{\omega}{2} = \dfrac{\alpha+\delta}{2}, \\ -\cos a + i\cos b = \dfrac{\gamma i}{\sin\frac{\omega}{2}}, & \\ & \sin\dfrac{\omega}{2} = \pm\sqrt{1-\left(\dfrac{\alpha+\delta}{2}\right)^2}. \\ -\cos c = \dfrac{(\alpha-\delta)i}{2\sin\frac{\omega}{2}}, & \end{cases}$$

Through these formulas the exercise posed above is solved: *the geometric elements of the rotation are expressed in terms of the parameters α, β, γ, δ.*

The solution of the opposite exercise: *to calculate the parameters α,*

β, γ, δ corresponding to a rotation given in terms of the axis and angle, is contained in the same equations. One finds immediately

$$(13') \qquad \begin{cases} \alpha = \cos\frac{\omega}{2} + i\sin\frac{\omega}{2}\,\cos c, \\ \beta = i\sin\frac{\omega}{2}\,(\cos a + i\cos b), \\ \gamma = i\sin\frac{\omega}{2}\,(\cos a - i\cos b), \\ \delta = \cos\frac{\omega}{2} - i\sin\frac{\omega}{2}\cos c. \end{cases}$$

These formulas show that the signs of the parameters α, β, γ, δ are now determined as well, since the parameters depend only on $\cos\frac{\omega}{2}$ and the products $\sin\frac{\omega}{2}\cos a$, etc.

The same holds, naturally, for the quaternion quantities A, B, C, D, for which the corresponding formulas become especially simple; they are, namely,

$$(14) \qquad \begin{cases} A = \sin\frac{\omega}{2}\cos a, \\ B = \sin\frac{\omega}{2}\cos b, \\ C = \sin\frac{\omega}{2}\cos c, \\ D = \cos\frac{\omega}{2}. \end{cases}$$

Finally, we return once again to the composition formulas (4) on page 32. We imagine the parameters α, β, γ, δ and α', β', γ', δ' in those formulas to be given uniquely, in the sense that we fix the corresponding rotation angles ω and ω' modulo 4π. We will presume, in the first place, that a corresponding convention is also necessary with respect to the resulting rotation, so that the parameters α'', β'', γ'', δ'' are again determined at first only up to the sign. Our formulas, however, give uniquely determined values of α'', β'', γ'', δ'' for specified values of $\alpha, \ldots$ and $\alpha', \ldots$; a specific choice between the two possible signs of $\alpha'', \ldots$ is therefore made by our formulas in advance.

We determine the principle through which this choice is made by comparing our current formulas with the geometric construction for the resulting rotation, which is referred to on page 10 of the first section. That construction and these formulas necessarily give the same value of the resulting rotation angle ω'' modulo 2π; it is at first questionable,

in contrast, whether the angle $\dfrac{\omega''}{2}$ in the two cases is equal modulo 2π or differs by π.

According to the geometric construction on page 10, the angle $\dfrac{\omega''}{2}$ is given as the exterior angle of a spherical triangle in which the opposite angles are $\dfrac{\omega}{2}$ and $\dfrac{\omega'}{2}$, respectively, and the opposite side is equal to the angle $(\widehat{12})$ that the axis of the first rotation encloses with the axis of the second.

On the other hand, our composition formulas give (one begins most conveniently from the last of equations (5)), with consideration of equations (14), the value

$$\begin{aligned}\cos\frac{\omega''}{2} &= \cos\frac{\omega}{2}\cos\frac{\omega'}{2} - \sin\frac{\omega}{2}\sin\frac{\omega'}{2}(\cos a\cos a' + \cos b\cos b' + \cos c\cos c') \\ &= \cos\frac{\omega}{2}\cos\frac{\omega'}{2} - \sin\frac{\omega}{2}\sin\frac{\omega'}{2}\cos(\widehat{12}).\end{aligned}$$

This is, however, one of the well-known fundamental formulas of spherical trigonometry. It shows us that the thus determined value of $\dfrac{\omega''}{2}$ can once again be interpreted as the exterior angle of a spherical triangle whose opposite angles are $\dfrac{\omega}{2}$ and $\dfrac{\omega'}{2}$ and that this value therefore coincides with the geometrically determined value of $\dfrac{\omega''}{2}$. Thus we can say:

The composition formulas (4) *and* (5) *are simply the analytic expression of the construction cited in the first section, since they give not only the same value of* ω'', *but also the same value of* $\dfrac{\omega''}{2}$.

§5. Passage to the so-called infinitesimal rotations

We now make the *limit passage from a finite to an "infinitesimal rotation,"* which has been mentioned already in §1, and therefore let the rotation angle ω decrease without bound, while we assume that the rotational velocity, which may be denoted by

$$\Omega = \frac{\omega}{dt},$$

approaches a finite limit. We operate with an infinitesimal rotation just as we are accustomed to operate in the infinitesimal calculus with

an "infinitesimal displacement," in that we neglect higher powers compared to lower powers in the limit of vanishing magnitude.

A single infinitesimal rotation may be supposed first; that is, the x, y, z and X, Y, Z systems are to be related directly by an infinitesimal rotation.

If we take up the concept of the rotation vector on page 11, we separate the angular velocity lying on the rotation axis into components with respect to the coordinate axes, and call these components p, q, r. We then have, evidently,

$$\Omega = \sqrt{p^2 + q^2 + r^2}, \quad \omega = \Omega\, dt,$$

and, with the use of the direction cosines of the rotation axis,

$$\begin{aligned} p &= \Omega \cos a, \\ q &= \Omega \cos b, \\ r &= \Omega \cos c. \end{aligned}$$

We ask, further, how the parameters α, β, γ, δ or A, B, C, D of the infinitesimal rotation are expressed in terms of p, q, r. The expressions for the quaternion quantities A, B, C, D are particularly simple. According to equations (14) of the previous section, we have, in the limit $\omega = 0$,

$$(1) \qquad \begin{cases} A = \frac{1}{2}\,\omega \cos a = \frac{1}{2}\, p\, dt, \\ B = \frac{1}{2}\,\omega \cos b = \frac{1}{2}\, q\, dt, \\ C = \frac{1}{2}\,\omega \cos c = \frac{1}{2}\, r\, dt, \\ D = 1. \end{cases}$$

Thus the expressions for our parameters α, β, γ, δ are

$$(2) \qquad \begin{cases} \alpha = D + iC = 1 + \dfrac{ir}{2}\, dt, & \beta = -\,B + iA = \dfrac{ip - q}{2}\, dt, \\ \gamma = B + iA = \dfrac{ip + q}{2}\, dt, & \delta = \quad D - iC = 1 - \dfrac{ir}{2}\, dt. \end{cases}$$

We also wish to write the explicit transformation equations for the rectangular coordinates xyz and XYZ in the case of an infinitesimal rotation. For this purpose, we need only place the above values of A, B, C, D into the schema (10) on page 22. We thus find, if we consistently reject the higher powers of dt, the skew determinant schema

(3)

	X	Y	Z
x	1	$-r\,dt$	$+q\,dt$
y	$+r\,dt$	1	$-p\,dt$
z	$-q\,dt$	$+p\,dt$	1

We may also give the equations of this schema in a somewhat different form. The coordinates of a fixed point in the top before the rotation are XYZ, and the coordinates of the same point after the rotation are xyz, both sets of coordinates referring to a coordinate system fixed in space. Thus

$$x - X, \quad y - Y, \quad z - Z$$

are the components of the displacement of the point; the velocity components x', y', z' are calculated from these displacements by the equations

$$x'\,dt = x - X, \quad y'\,dt = y - Y, \quad z'\,dt = z - Z.$$

On the basis of the schema (3), *the linear velocity of a point of the top thus becomes*

$$(3') \qquad \begin{cases} x' = - rY + qZ, \\ y' = rX - pZ, \\ z' = -qX + pY. \end{cases}$$

We may write x, y, z on the right-hand side of these equations instead of X, Y, Z, since these quantities differ only by terms with the factor dt. We thus have

$$(3'') \qquad \begin{cases} x' = - ry + qz, \\ y' = rx - pz, \\ z' = -qx + py. \end{cases}$$

We can now verify the important theorem on the composition of two infinitesimal rotations that was previously assumed on page 11, and that justifies the introduction of the words “rotation vector.” We consider, in addition to the rotation with the parameters α, β, γ, δ in equations (2), a second rotation that is given by the parameters

$$\alpha' = 1 + \frac{ir'}{2}\,dt, \qquad \beta' = \frac{ip' - q'}{2}\,dt,$$
$$\gamma' = \frac{ip' + q'}{2}\,dt, \qquad \delta' = 1 - \frac{ir'}{2}\,dt.$$

According to the composition rule of page 32, the parameters of the

resultant rotation are, neglecting the higher powers of dt,

$$\alpha'' = 1 + \frac{i(r+r')}{2}\,dt, \qquad \beta'' = \frac{i(p+p')-(q+q')}{2}\,dt,$$
$$\gamma'' = \frac{i(p+p')+(q+q')}{2}\,dt, \qquad \delta'' = 1 - \frac{i(r+r')}{2}\,dt.$$

The components of the resulting rotational velocity are thus

$$p'' = p + p', \quad q'' = q + q', \quad r'' = r + r'.$$

Two infinitesimal rotations therefore compose as vectors; namely, their components simply add.

In particular, the resultant of two infinitesimal rotations is independent of their order; *infinitesimal rotations*, as we said, *are interchangeable operations.*

We also now see clearly the general basis of this simple result. It lies essentially in the fact that we are allowed to neglect the higher powers of dt in the limit of vanishingly small quantities, and that in the remaining first powers the quantities $p, \ldots, p', \ldots$ necessarily appear in the combinations $p+p', \ldots$. One immediately recognizes here that a corresponding result must occur for any "infinitesimal transformation."

We could, moreover, have derived the very well known equations contained in the schema (3) more simply, with the help of this last theorem, by composing, as in fact is usually done, the successive rotations $p\,dt$, $q\,dt$, $r\,dt$.

We next combine a finite rotation with an infinitesimal rotation. The finite rotation has parameters α, β, γ, δ, and may be applied first. The parameters of the infinitesimal rotation are α', β', γ', δ'. We ask for the parameters

$$\alpha'' = \alpha + d\alpha, \quad \beta'' = \beta + d\beta, \quad \gamma'' = \gamma + d\gamma, \quad \delta'' = \delta + d\delta$$

of the resulting rotation. Since we will now assume that the moving coordinate frame XYZ coincides with the fixed frame xyz in its initial position *before the first rotation*, its position *at the beginning of the second* (infinitesimal) *rotation* differs from that of the fixed frame. From now on, therefore, we must carefully distinguish the components of the rotation vector with respect to the moving frame and the components of the rotation vector with respect to the fixed frame. We call the former set of components p, q, r, and the latter set π, κ, ϱ. If we wish to apply the composition formulas of page 32 directly, we must refer the infinitesimal rotation to the moving frame, and thus use

the components p, q, r. We obtain, with consideration of equations (2),

$$\alpha + d\alpha = \alpha\alpha' + \beta\gamma' = \alpha + \left(\frac{ir}{2}\,\alpha + \frac{ip+q}{2}\,\beta\right) dt,$$

$$\beta + d\beta = \alpha\beta' + \beta\delta' = \beta + \left(\frac{ip-q}{2}\,\alpha - \frac{ir}{2}\,\beta\right) dt, \quad \text{etc.}$$

Thus we write, adjoining the analogous equations for γ and δ,

$$(4) \qquad \begin{cases} \dfrac{d\alpha}{dt} = \dfrac{ir}{2}\,\alpha + \dfrac{ip+q}{2}\,\beta, & \dfrac{d\beta}{dt} = \dfrac{ip-q}{2}\,\alpha - \dfrac{ir}{2}\,\beta, \\ \dfrac{d\gamma}{dt} = \dfrac{ir}{2}\,\gamma + \dfrac{ip+q}{2}\,\delta, & \dfrac{d\delta}{dt} = \dfrac{ip-q}{2}\,\gamma - \dfrac{ir}{2}\delta. \end{cases}$$

The preceding equations give the changes in the parameters α, β, γ, δ due to the application of an infinitesimal rotation with components p, q, r. If we imagine the components p, q, r to be given somehow as functions of time, then we can determine from them the successive changes in α, β, γ, δ, and thus the successive changes in position of the top in space, through the integration of this set of differential equations. If, conversely, the parameters α, β, γ, δ are known as functions of time, then the successive positions of the top in space are given, and we can compute the components p, q, r from our equations by differentiation.

Now the ratios $p:q:r$ determine the position of the instantaneous axis of rotation with respect to the frame XYZ fixed in the top, and the successive values of these ratios determine the *form of the polhode cone*. Further, the components p, q, r denote the coordinates of the endpoint of the rotation vector in the same coordinate system; the successive values of p, q, r themselves thus give the *form of the polhode curve*.

We find the explicit expressions for p, q, r by the solution of equations (4). For simplicity, we retain the complex combinations $p + iq$, $-p + iq$, $-r$, and obtain

$$(5) \quad \begin{cases} p + iq = 2i\left(\beta\,\dfrac{d\delta}{dt} - \delta\,\dfrac{d\beta}{dt}\right), \\ -p + iq = 2i\left(\alpha\,\dfrac{d\gamma}{dt} - \gamma\,\dfrac{d\alpha}{dt}\right), \\ -r \qquad\;\; = 2i\left(-\alpha\,\dfrac{d\delta}{dt} + \gamma\,\dfrac{d\beta}{dt}\right) = 2i\left(-\beta\,\dfrac{d\gamma}{dt} + \delta\,\dfrac{d\alpha}{dt}\right), \end{cases}$$

where the two given values for $-r$ are identical because of the relation $\alpha\delta - \beta\gamma = 1$.

We attach here the formulas for the *curve* and *cone of the herpolhode*, thus computing the components π, κ, ϱ of the rotation vector with

respect to the fixed frame. We can proceed by placing the previously determined values of $p+iq$, $-p+iq$, $-r$ in the schema (9) of page 21 instead of Ξ, H, Z; the corresponding values of ξ, η, ζ then yield the desired quantities $\pi+i\kappa$, $-\pi+i\kappa$, $-\varrho$. However, the following path leads more directly to the goal. We remarked on page 14 that the herpolhode curve of the direct motion lies diametrically opposed, with respect to the support point, to the polhode curve of the inverse motion. Now we obtain the polhode curve of the inverse motion from equations (5) if we replace, according to the rule given on page 30, the parameters α, β, γ, δ by δ, $-\beta$, $-\gamma$, α. Thus the *formulas for the herpolhode curve* of the direct motion are

$$(6)\qquad \begin{cases} \pi+i\kappa = 2i\left(\beta\,\dfrac{d\alpha}{dt} - \alpha\,\dfrac{d\beta}{dt}\right), \\ -\pi+i\kappa = 2i\left(\delta\,\dfrac{d\gamma}{dt} - \gamma\,\dfrac{d\delta}{dt}\right), \\ -\varrho \qquad = 2i\left(\delta\,\dfrac{d\alpha}{dt} - \gamma\,\dfrac{d\beta}{dt}\right) = 2i\left(\beta\,\dfrac{d\gamma}{dt} - \alpha\,\dfrac{d\delta}{dt}\right). \end{cases}$$

The *herpolhode cone* is naturally determined by the ratios of the quantities on the left-hand side.

We do not wish to write the corresponding equations for A, B, C, D in detail; they follow from the separation of (5) and (6) into real and imaginary parts. We will later use, however, the representations for p, q, r in terms of the Euler angles φ, ψ, ϑ and their derivatives with respect to time. We obtain these representations, for example, by starting from equations (5). We compute the combinations of α, β, γ, δ on the right-hand sides with the help of the original definitions of our parameters in terms of φ, ψ, ϑ. We thus obtain, for example,

$$\begin{aligned} \beta\,\frac{d\delta}{dt} - \delta\,\frac{d\beta}{dt} &= -\,\frac{i}{2}\,e^{-i\varphi}\,\vartheta'\left(\sin^2\frac{\vartheta}{2} + \cos^2\frac{\vartheta}{2}\right) + e^{-i\varphi}\,\psi'\sin\frac{\vartheta}{2}\cos\frac{\vartheta}{2} \\ &= -\,\frac{i}{2}\,(\vartheta' + i\psi'\sin\vartheta)e^{-i\varphi} \end{aligned}$$

and

$$\begin{aligned} -\,\alpha\,\frac{d\delta}{dt} + \gamma\,\frac{d\beta}{dt} &= +\,\frac{i}{2}\cos^2\frac{\vartheta}{2}\;(\varphi' + \psi') - \frac{i}{2}\sin^2\frac{\vartheta}{2}\;(-\,\varphi' + \psi') \\ &= \frac{i}{2}\,(\varphi' + \cos\vartheta\,\psi'); \end{aligned}$$

thus

$$\begin{aligned} p+iq &= (\;\;\vartheta' + i\psi'\sin\vartheta)e^{-i\varphi}, \\ -\,p+iq &= (-\vartheta' + i\psi'\sin\vartheta)e^{+i\varphi}, \\ -\,r \qquad &= -\,(\varphi' + \psi'\cos\vartheta), \end{aligned}$$

or

$$(7)\qquad \begin{cases} p = \vartheta' \cos\varphi + \psi' \sin\vartheta \sin\varphi, \\ q = -\,\vartheta' \sin\varphi + \psi' \sin\vartheta \cos\varphi, \\ r = \varphi' + \cos\vartheta\, \psi'. \end{cases}$$

The passage to the inverse motion, which, according to page 31, is accomplished in terms of the Euler angles by the interchange of φ, ψ, ϑ with $-\psi$, $-\varphi$, $-\vartheta$, yields

$$(8)\qquad \begin{cases} \pi = \vartheta' \cos\psi + \varphi' \sin\vartheta \sin\psi, \\ \kappa = \vartheta' \sin\psi - \varphi' \sin\vartheta \cos\psi, \\ \varrho = \psi' + \cos\vartheta\, \varphi'. \end{cases}$$

Solved for φ', ψ', ϑ', equations (7) are

$$(9)\qquad \begin{cases} \varphi' = r - \dfrac{\cos\vartheta}{\sin\vartheta}\,(p \sin\varphi + q \cos\varphi), \\ \psi' = \dfrac{1}{\sin\vartheta}\,(p \sin\varphi + q \cos\varphi), \\ \vartheta' = (p \cos\varphi - q \sin\varphi). \end{cases}$$

One can, naturally, also obtain equations (7) and (9) through the elementary separation of the infinitesimal rotation into components, as is often carried out in textbooks.

Equations (7) and (8) give a new representation for the cones of the polhode and herpolhode, which is certainly inferior to the previous in clarity; equations (9) state how the parameters φ, ψ, ϑ are modified by the addition of an infinitesimal rotation; they form one part of the system of differential equations whose integration will occupy us in the fourth chapter. —

In the second section, we presented different parameter systems (α, β, γ, δ; A, B, C, D; φ, ψ, ϑ) through which we determined the instantaneous *position* of the top; on the other hand, we have become acquainted in this section, as a result of the study of infinitesimal rotations, with parameter systems through which the instantaneous *state of velocity* of the top is characterized. These are, in the first place, the quantities p, q, r and φ', ψ', ϑ'. —

We can denote these quantities briefly as the *velocity coordinates of the top*, and, in particular, call p, q, r the *rectangular velocity coordinates*, and φ', ψ', ϑ' the *skew-angular velocity coordinates*, since, in the former case, we separate the rotation vector into components with respect to the rectangular axes X, Y, Z, and, in the latter case, we separate the rotation vector into components with respect to three

generally skew-angular axes (namely, the figure axis, the vertical, and the line of nodes).

There is a further important distinction between the triplets p, q, r and φ', ψ', ϑ': *the quantities* φ', ψ', ϑ' *are the time derivatives of certain spatial (angular) measures, while the quantities* p, q, r *are not*; or, *the quantities* $\varphi'\,dt$, $\psi'\,dt$, $\vartheta'\,dt$ *are exact differentials, while the quantities* $p\,dt$, $q\,dt$, $r\,dt$ *are inexact differentials.*

We recognize this, for example, from the fact that the values of the time integrals

$$\int p\,dt, \quad \int q\,dt, \quad \int r\,dt$$

from an initial time t_0 to a final time t_1 depend not only on the initial and final positions of the body, but also on the intermediate positions, which is evident geometrically. The same conclusion follows from the consideration of any one of the latterly written equations. We take, for example, the first of equations (7):

$$p\,dt = d\vartheta\,\cos\varphi + d\psi\,\sin\vartheta\sin\varphi.$$

If the right-hand side is to be a perfect differential, then

$$\frac{\partial\cos\varphi}{\partial\psi} \quad \text{and} \quad \frac{\partial\sin\vartheta\sin\varphi}{\partial\vartheta}$$

must have the same value, which is evidently not the case.

We must therefore say: *the quantities* p, q, r *are not, as are the velocity components* x', y', z' *of a single point mass, time derivatives of spatial measures, but only linear functions of such time derivatives.* Nevertheless, they are very appropriate as velocity coordinates in the theory of the top, and are also generally common.

We can, naturally, characterize the state of velocity of the top through arbitrarily many other more or less appropriate parameters; for example, as previously, through the quantities $\frac{d\alpha}{dt}, \frac{d\beta}{dt}, \frac{d\gamma}{dt}, \frac{d\delta}{dt}$. These latter are obviously "supernumerary velocity coordinates," while our quantities p, q, r or φ', ψ', ϑ' are to be designated as "independent velocity coordinates."

At this point we can immediately give the velocity coordinates of the *freely moving rigid body.* For this purpose, we have only to first establish, on the one hand, the velocity of the reference point (the translational velocity), and, on the other hand, the motion about the reference point (the rotational velocity), each through three parameters. The simplest and generally used velocity coordinates of the freely moving rigid body are the six quantities

$$x',\ y',\ z',\ p,\ q,\ r,$$

where x, y, z denote, as at the end of §2, the rectangular coordinates of the reference point. We can naturally, however, characterize the state of velocity in a great variety of ways through six other parameters. Each such set of six quantities may be conceived as the "*coordinates of the instantaneous motion-screw*" through which we can characterize, according to §1, any infinitesimal motion of a free rigid body.

§6. The example of regular precession

As an example of the preceding general theory, we consider *regular precession*, a particularly simple motion of the top that we shall call upon repeatedly for orientation in the following study of more complicated motions. Naturally, only the *kinematic* side of this motion can be considered here; its *kinetic* investigation (that is, the answers to the questions of whether and under what circumstances a regular precession can be attained for a given top) will occupy us later. We wish to conduct the investigation first geometrically, in accordance with the first section of this chapter, and then analytically, in accordance with the previous section.

We define regular precession by giving, on the one hand, the motion of the figure axis in space, and, on the other hand, the motion of the top with respect to the figure axis. Both components of the motion are, in the present case, as simple as possible. *Namely, the figure axis rotates about a fixed line in space with constant inclination and uniform angular velocity; at the same time, the top turns about the figure axis, likewise with uniform angular velocity.* We denote the fixed line in space as the "axis of precession"; we usually imagine this axis, for the sake of simplicity, to be vertical.

The rotational velocity of the figure axis about the axis of precession is called ν, and the additional rotational velocity of the top about the figure axis is called μ; we designate the former as the "precessional velocity of the figure axis." We first wish to assume, moreover, so as to avoid the consideration of differing cases as much as possible in the preliminary geometric analysis, that the rotational velocity μ is very large in relation to the "precessional velocity" ν, as will in fact be the rule in later applications.

Our individual rotations μ and ν combine at each moment into a resultant rotation that determines, through its angular velocity, axis, and sense, the magnitude, direction, and sense of the instantaneous rotation vector. In the figures below, OV denotes the axis of precession and OF

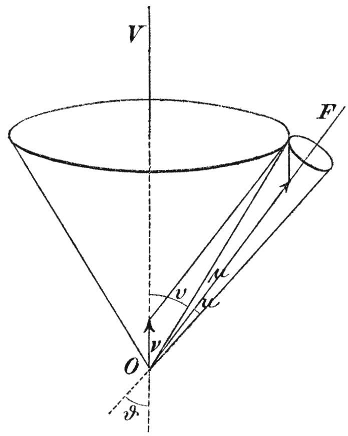


Fig. 5.

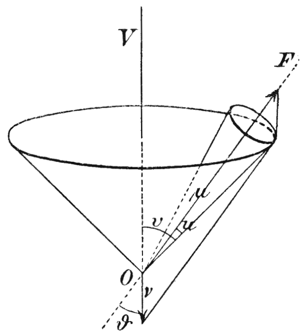


Fig. 6.

denotes the figure axis. We can choose the designation “figure axis” so that the angle between OV and OF is not greater than a right angle, and we wish to exclude, for the present, the special case in which this angle is exactly equal to a right angle. On the figure axis and the axis of precession, we mark off the angular velocities μ and ν, respectively, in the manner that is established on page 11. We distinguish two cases, according to whether the angular velocities μ and ν have the same or the opposite sense, and therefore whether the corresponding vectors enclose an acute or an obtuse angle. We designate the precession in the first case as *progressive*, and in the second case as *retrograde*. The diagonal of the parallelogram construction lies, in the first case, in the acute angle between the vertical and the figure axis, and, in the second case, outside this angle. Moreover, our diagonal has a fixed length and an unchanging relative position with respect to the vertical and to the figure axis in both cases, since by assumption the component vectors μ and ν have constant lengths and unchanging relative positions.

During the course of the motion, the endpoint of the rotation vector describes, in so far as we consider its position in space, a circle about the vertical; at the same time it traverses, in so far as we consider its position with respect to the top, a circle about the figure axis. *In the case of regular precession, the curves of the polhode and the herpolhode are thus simple circles; correspondingly, the cones of the polhode and*

the herpolhode, which those curves project from O, become ordinary circular cones. With respect to the relative position of these two cones, we confirm on the basis of Figs. 5 and 6 a distinction: *the polhode cone rolls on the outside of the herpolhode cone in progressive precession, while in retrograde precession the polhode cone rolls on the inside of the herpolhode cone.*

We may derive some additional elementary geometric consequences from the parallelogram construction. We denote, as earlier, the angle between the figure axis and the vertical by ϑ. Further, let the angle between the instantaneous rotation axis and the figure axis be u, and the angle between the instantaneous rotation axis and the vertical be v. Then u determines the opening of the polhode cone, and v determines the opening of the herpolhode cone.

If we denote, as earlier, the magnitude of the resultant rotational velocity by Ω, which is therefore the length of the diagonal of the parallelogram, then we have, according to the Pythagorean theorem,

$$\Omega^2 = \mu^2 + \nu^2 + 2\mu\nu\cos\vartheta.$$

Further, it follows from the two triangles into which our parallelogram is divided by the diagonal that

$$\frac{\sin u}{\nu} = \frac{\sin\vartheta}{\Omega}$$

and

$$\frac{\sin v}{\mu} = \frac{\sin\vartheta}{\Omega}.$$

Therefore

(1) $$\mu\sin u = \nu\sin v.$$

An important example of retrograde precession is given by our *Earth.* The Earth plays here the role of the top; we imagine the midpoint of the Earth to be fixed during the motion. The vertical corresponds to the normal to the plane of the ecliptic. *The axis of the Earth circles this line with a fixed inclination of approximately* $23\,{}^1\!/_2{}^\circ$ *in approximately* $26\,000$ *years.* If we take the length of a day as the time unit, then

$$\mu = 2\pi, \qquad \nu = \frac{-2\pi}{365\,.\,26\,000}.$$

From equation (1) there follows

(2) $$\sin u = \frac{-\sin v}{365\,.\,26\,000}.$$

The angle u is therefore a very small quantity, so that we can replace

$\sin u$ by u. Since, further, as in Fig. 6, $v = \vartheta + u$, we may replace v by ϑ. *The herpolhode cone is thus very nearly a circular cone with an opening angle of* $23\,{}^1\!/_2{}^\circ$. We ask, further, for the form of the polhode cone. Instead of its opening angle, we prefer to give the radius (r) of the circle formed by the intersection of this cone with the surface of the Earth. If we denote the radius of the Earth by R*) $\left(R = \frac{40\,000\,000}{2\pi} \text{ Meter} = \frac{4\,000\,000\,000}{2\pi} \text{ cm}\right)$, then we have, according to equation (2),

$$r = Ru = \frac{\sin(23\,{}^1\!/_2{}^\circ)\,.\,4\,000\,000\,000}{2\pi\,.\,365\,.\,26\,000}\text{ cm} = 27\text{ cm}.$$

The trace of the cone of the polhode on the surface of the Earth would therefore be a small circle of radius 27 *cm inscribed around the North Pole.* We can thus imagine the precessional motion of the Earth as that of a correspondingly thin polhode cone which rolls on the interior of a herpolhode cone with an opening angle of approximately $23\,{}^1\!/_2{}^\circ$.

As we see, the Poinsot theory of the rolling cones gives, in the case of regular precession, a highly intuitive and concise image of the course of the motion. Poinsot himself treats the example of regular precession with special predilection in the applications of his theory**).

We now consider regular precession once again analytically, on the basis of the results derived in the previous section. We first carry over the geometric definition of our motion into analysis, which is achieved most simply with the help of the Euler parameters φ, ψ, ϑ.

Namely, we obviously have

$$\vartheta = \text{const.}, \quad \varphi = \mu t, \quad \psi = \nu t. \tag{3}$$

From equations (7) of the previous section, the components of the instantaneous rotation vector become, if we refer them to the coordinate frame fixed in the top,

$$p = \nu \sin\vartheta \sin\varphi, \quad q = \nu \sin\vartheta \cos\varphi, \quad r = \mu + \nu\cos\vartheta, \tag{4}$$

and, according to equations (8) of the same section, if we calculate them for the coordinate frame fixed in space,

$$\pi = \mu \sin\vartheta \sin\psi, \quad \kappa = -\mu \sin\vartheta \cos\psi, \quad \varrho = \nu + \mu\cos\vartheta. \tag{5}$$

These equations again show *that the curves of the polhode and the herpolhode become circles in our case.* In fact we have, for example,

$$p^2 + q^2 = \nu^2 \sin^2\vartheta = \text{const.}, \quad r = \mu + \nu\cos\vartheta = \text{const.}$$

*) All calculated in round numbers.

**) Cf., for example, his work: Théorie des cônes circulaires roulants, in the journal Connaissance des temps, 1853; this work also appears in Liouville's Journal, ser. 1, t. 18, 1853, p. 41.

If u and v have the meanings given above, the opening angles of the rolling cones are calculated according to the formulas

$$\operatorname{ctg} u = \frac{r}{\sqrt{p^2+q^2}}, \quad \operatorname{ctg} v = \frac{\varrho}{\sqrt{\pi^2+\kappa^2}}; \tag{6}$$

we therefore have, from equations (4) and (5),

$$\operatorname{ctg} u = \frac{\frac{\mu}{\nu}+\cos\vartheta}{\sin\vartheta}, \quad \operatorname{ctg} v = \frac{\frac{\nu}{\mu}+\cos\vartheta}{\sin\vartheta}. \tag{7}$$

These values of u and v naturally satisfy the above relation (1), as one easily calculates.

We now give a *detailed discussion of the different possible cases of regular precession motion*, which we classify according to the value of $\frac{\nu}{\mu}$. The present classification subsumes the previous, in which we distinguished the motion as progressive or retrograde, according to whether $\frac{\nu}{\mu}$ was greater than or less than zero.

In the following discussion, we imagine ϑ as fixed, and we can assume, as above, that ϑ is not greater than a right angle. If we postpone the limiting cases $\vartheta = 0$ and $\vartheta = \frac{\pi}{2}$ until the end, we assume, for the present,

$$0 < \vartheta < \frac{\pi}{2}.$$

We may take one of the opening angles u and v, say v, as an acute angle. Namely, at the same time that the cones u and v roll on one another, the diametral cones $\pi - u$ and $\pi - v$ roll on one another as well. Thus if $v > \frac{\pi}{2}$, we can attain $v' < \frac{\pi}{2}$ through the passage to the cones $u' = \pi - u$ and $v' = \pi - v$.

The entire range of values of $\frac{\nu}{\mu}$ is divided, for our purpose, into four intervals, which are separated by the four limiting values

$$\frac{\nu}{\mu} = \infty, \quad \frac{\nu}{\mu} = 0, \quad \frac{\nu}{\mu} = -\cos\vartheta, \quad \frac{\nu}{\mu} = -\frac{1}{\cos\vartheta}.$$

First limiting case: $\frac{\nu}{\mu} = +\infty$. In this case $\mu = 0$; that is, the top does not rotate relative to its figure axis at all. The formulas (7) show that $v = 0$ as well. The herpolhode cone is therefore infinitely thin, and the instantaneous axis of rotation always coincides with the vertical.

In our first limiting case, the top executes a uniform rotation with angular velocity ν about a fixed line (namely, the vertical) different from the figure axis.

First interval: $+\infty > \dfrac{\nu}{\mu} > 0$. In the interior of our first interval, ν and μ have the same signs. The parallelogram construction which we carried out above then shows that the instantaneous axis of rotation lies in the acute angle between the vertical and the figure axis. Correspondingly, it follows from equations (7) that u and v are smaller than ϑ. *The moving cone thus rolls on the outside of the fixed cone* (cf. Fig. 7). We will designate this motion as *epicycloidal.*

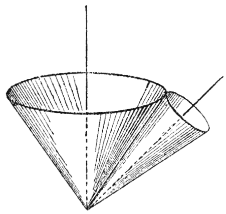


Fig. 7.

Second limiting case: $\dfrac{\nu}{\mu} = 0$. With decreasing positive values of $\dfrac{\nu}{\mu}$, the herpolhode cone gradually widens, while the polhode cone contracts. In the limiting case $\dfrac{\nu}{\mu} = 0$, the polhode cone has become infinitely thin; in fact, equations (7) give $u = 0$ in this case. Consequently, the instantaneous axis of rotation always coincides with the figure axis. At the same time, the figure axis remains fixed in space, because $\nu = 0$. *The top therefore rotates with constant angular velocity μ about its figure axis, which is fixed in space.*

Second interval: $0 > \dfrac{\nu}{\mu} > -\cos\vartheta$. If $\dfrac{\nu}{\mu}$ becomes negative, the polhode cone begins to widen again. At the same time, it appears in the interior of the herpolhode cone. In fact, it follows from equations (7) that $v > \vartheta$. *The moving cone thus rolls on the interior of the fixed cone* (cf. Fig. 8). We will designate this case as *hypocycloidal motion.*

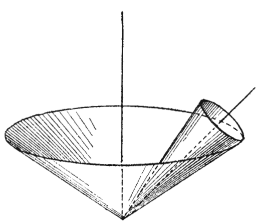


Fig. 8.

Third limiting case: $\dfrac{\nu}{\mu} = -\cos\vartheta$. According to equations (7), when $\dfrac{\nu}{\mu}$ has attained the value $-\cos\vartheta$, v has become equal to $\dfrac{\pi}{2}$, *and thus the herpolhode cone has degenerated into a plane.*

Third interval: $-\cos\vartheta > \dfrac{\nu}{\mu} > -\dfrac{1}{\cos\vartheta}$. If $\dfrac{\nu}{\mu}$ decreases fur-

ther, v becomes greater than $\frac{\pi}{2}$. According to our previous agreement, we must then go over to the diametral cones $\pi - u$ and $\pi - v$, through which $v < \frac{\pi}{2}$ and $u > \frac{\pi}{2}$. *In the third interval, an obtuse polhode cone thus rolls on the exterior of an acute herpolhode cone* (cf. Fig. 9). We wish to designate this type of motion as *anticycloidal.*

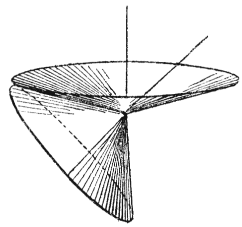


Fig. 9.

Fourth limiting case: $\frac{\nu}{\mu} = -\frac{1}{\cos\vartheta}$. As $\frac{\nu}{\mu}$ decreases from $-\cos\vartheta$ to $-\frac{1}{\cos\vartheta}$, the polhode cone gradually flattens. For the limiting value $\frac{\nu}{\mu} = -\frac{1}{\cos\vartheta}$, it degenerates, according to equations (7), *into a plane.*

Fourth interval: $-\frac{1}{\cos\vartheta} > \frac{\nu}{\mu} > -\infty$. With further decrease of $\frac{\nu}{\mu}$, the polhode cone contracts, so that in the fourth interval u is again less than $\frac{\pi}{2}$. *At the same time, the polhode cone encloses the cone of the herpolhode from the exterior* (cf. Fig. 10). We wish to designate the type of rolling here as *pericycloidal motion.*

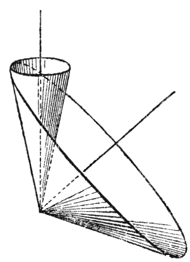


Fig. 10.

Finally, the limiting case $\frac{\nu}{\mu} = -\infty$ is identical with the limiting case $\frac{\nu}{\mu} = +\infty$ with which we began this discussion.

To place our current classification in relation to the previous, we adjoin the following schema, which will be understandable without further explanation:

Retrograde,			Progressive Precession
Peri-	Anti-	Hypo-	Epi-cycloidal Motion

$$-\infty \longleftarrow \quad -\frac{1}{\cos\vartheta} \quad | \quad -\cos\vartheta \quad | \quad 0 \quad | \quad \longrightarrow +\infty$$

It remains only to say a word concerning the limiting cases $\vartheta = 0$ and $\vartheta = \frac{\pi}{2}$.

If $\vartheta = 0$, the figure axis coincides with the vertical, has a fixed direction in space, and is at the same time the rotation axis. *The top rotates with constant velocity about this axis; the rolling cones are infinitely thin.* The position of the line of nodes in the equatorial plane is undetermined, as are the values of ν and μ.

In the other limiting case $\vartheta = \frac{\pi}{2}$, the figure axis describes the horizontal plane. The second and fourth of our above intervals are eliminated; the motion is either *epi-* or *anticycloidal.* —

We have taken up regular precession all the more willingly because, allowing for a small modification of the concept, any arbitrary motion of the top at a specified point in time can be conceived as an appropriate precessional motion. The modification of the concept is that we go over from a regular (or uniform) precession to a *uniformly accelerated precession.* We understand by this a motion that appears just as a regular precession with respect to the successive directions of the rotation axis, but that differs from a regular precession with respect to the magnitude of the rotation, in that this magnitude increases in proportion with time. The cones of the polhode and herpolhode in this motion are again simple circular cones; the corresponding curves, however, are not circles, but rather spiral-shaped lines that wander around the circular cones in ever widening windings.[35]

Now let an arbitrary motion be given through its polhode and herpolhode cones and curves. We have seen, in the first section, that every motion of the top can be replaced at a single instant by a *simple rotation* about the instantaneous axis of rotation. *This replaces the given motion with respect to both the direction and magnitude of the velocity of all the individual mass points of the top at a given point of time; or, as we can say, it approximates the motion to the first order.*

In a corresponding manner, we claim that *with a uniformly accelerated precession, we can approximate any arbitrary motion of the top to the second order with respect to the direction and magnitude of the motion; that is, approximate the motion at two neighboring moments of time.*

To see this, we construct the circular cones that osculate to the given cones of the polhode and herpolhode along the length of the instantaneous rotation axis. We can designate these as *curvature cones*, since they play exactly the same role for us as the curvature circles in the theory of curves. We consider, further, the polhode and herpolhode

curves of the given motion. These curves determine, by means of the direction in which they pass through the endpoint of the instantaneous rotation vector, a specified change in the length of the rotation vector at the considered point of time. We now construct on the curvature cones new polhode and herpolhode curves that exhibit the same constant change in velocity. The resulting spiral curves are tangent to the polhode and herpolhode curves of the given motion at the considered point of time. If we now roll the two curvature cones on each other, so that the rotational velocity at each moment is proportional to the length of the rotation vector determined by our polhode and herpolhode curves, then we obtain a uniformly accelerated precession that approximates the given motion to the second order with respect to both the direction and the magnitude of the velocity of all points of the top. In fact, the position and magnitude of the rotation vectors of the given motion and our uniformly accelerated precession agree at two neighboring moments of time.

If we were to reproduce only the *direction* of the velocity of the actual motion, thus disregarding the *magnitude* of the velocity, then we could have managed with a *regular* precession, which is produced by any *uniform* rolling of our curvature cones.

§7. Excursus on the theory of quaternions

We cannot conclude this chapter without pointing out the relation in which the previous discussions stand with respect to Hamilton's theory of quaternions. Although this theory is no longer young*), opinions of its worth vary strongly even today. The basis for this may be sought in the fact that the advocates of the quaternion theory generally present their work one-sidedly and burden it with a touch of metaphysics,[36] so that the simple geometric interpretation which can be given to the operations of their calculations does not always stand out in sufficient relief. In the following presentation, which is adapted freely to the preceding, we hope to provide, in this respect, a contribution to the clarification of the view; we do so all the more willingly since the theory of quaternions, as we will see below, contains the so-called *vector calculus* as a special case, and since this vector calculus has

*) Hamilton's fundamental work Lectures on Quaternions originated in the year 1853; the Elements of Quaternions (German by Glan) followed in 1866.

become a convenient and generally common means of expression for many problems of modern physics, of which we ourselves will make frequent use. —

We have previously designated four parameters A, B, C, D as "quaternion quantities," which we introduced for the analytic representation of rotations about a fixed point. Those parameters are a special case of the quantities that we will now designate by A, B, C, D. To arrive at these latter quantities, we make the following deliberation.

We consider *an operation that combines a rotation about O and a similarity transformation with the center O.* The rotation is determined through the direction of the rotation axis and the magnitude of the rotation angle, and may, as previously (cf. page 36), be characterized by the angles a, b, c, and $\frac{\omega}{2}$; the similarity transformation causes a certain enlargement of proportions, which we wish to denote by T. We propose the abbreviated expression *rotation-stretch*[37] for this combined operation. In our rotation-stretch, each point XYZ will be transformed into a point xyz. The relation between the coordinates of the two points is entirely similar to the corresponding relation in the case of a pure rotation. We have only, after we reach a point $x'y'z'$ from XYZ through a *rotation*, to multiply the coordinates by the enlarging proportionality T to arrive at the point xyz that represents the result of the *rotation-stretch*. Consequently, the factor T is added to each coefficient in the rotation schema of page 22. We can also, however, view that schema directly as the expression of a rotation-stretch if we slightly alter the meanings, as opposed to the previous definitions, of the parameters A, B, C, D. *Namely, we now define* (in contrast to equations (14) on page 38) A, B, C, D *through the expressions*

$$
\begin{aligned}
A &= \sqrt{T}\sin\frac{\omega}{2}\,\cos a, \\
B &= \sqrt{T}\sin\frac{\omega}{2}\,\cos b, \\
C &= \sqrt{T}\sin\frac{\omega}{2}\,\cos c, \\
D &= \sqrt{T}\cos\frac{\omega}{2}.
\end{aligned}
\tag{1}
$$

The transformation formulas of the rotation-stretch are thus identical to the previously written schema

(2)

	X	Y	Z
x	$D^2+A^2-B^2-C^2$	$2(AB-CD)$	$2(AC+BD)$
y	$2(AB+CD)$	$D^2-A^2+B^2-C^2$	$2(BC-AD)$
z	$2(AC-BD)$	$2(BC+AD)$	$D^2-A^2-B^2+C^2$

The difference between the current and the previous definitions of A, B, C, D is that these quantities are now *freely changeable parameters*, while they were previously bound by the relation

$$A^2 + B^2 + C^2 + D^2 = 1 \tag{3}$$

of page 21. We can say: *to each arbitrarily given set of values A, B, C, D there now corresponds a determined rotation-stretch.* Namely, in consequence of the defining equations (1), we calculate from the absolute values of A, B, C, D the enlarging proportionality of the rotation-stretch as

$$A^2 + B^2 + C^2 + D^2 = T, \tag{4}$$

while the axis and rotation angle of the rotation-stretch are given uniquely by the ratios $A : B : C : D$.

Conversely, there corresponds to each rotation-stretch an entirely determined parameter system A, B, C, D, assuming that we consider the rotation angle ω as given not only simply (that is, modulo 2π), but rather that we also fix, as previously agreed, $\frac{\omega}{2}$ in a specified way modulo 2π.

We next ask for the formula for the *composition of two rotation-stretches.* Let the first rotation-stretch be given by the parameters A, B, C, D, and the second by A', B', C', D'. The resultant of the two will be designated by A'', B'', C'', D''. We obviously arrive at the resulting rotation-stretch if we combine the corresponding rotations, on the one hand, and the stretches, on the other hand, individually. We have written the formulas for the composition of two rotations in terms of the quaternion quantities on page 33. The composition of two stretches is accomplished according to the simple formula

$$T'' = TT'.$$

Thus the composition formulas for rotation-stretches, with the new meanings of the parameters A, B, C, D, are just the same, formally, as those for rotations. We have

$$
(5) \qquad \begin{cases} A'' = AD' + BC' - CB' + DA', \\ B'' = - AC' + BD' + CA' + DB', \\ C'' = AB' - BA' + CD' + DC', \\ D'' = DD' - AA' - BB' - CC'. \end{cases}
$$

We now consider these formulas in the light of the theory of quaternions. We give the primitive definition of the word quaternion on the basis of our concept of the rotation-stretch: *a quaternion means nothing other than the operation of a rotation-stretch.* It is uniquely determined by the magnitude of the stretch (T), the axis of the rotation (a, b, c), and the magnitude of the *half* rotation angle $\left(\frac{\omega}{2}\right)$.

The quaternion thus receives four "components" A, B, C, D. To allow these components to appear from the exterior as belonging together, we join them in a symbolic expression by means of "three imaginary units" i, j, k. We will thus write a quaternion as a "four-term complex number"

$$
(6) \qquad Q = iA + jB + kC + D.
$$

The introduction of the imaginary units is, from our standpoint, purely conventional; it merely serves to allow the parameters A, B, C, D to appear, to a certain degree, as named quantities. In their original introduction by H a m i l t o n, the situation was different. Hamilton set out to generalize the usual complex numbers. The above notation was thus somewhat more essential for him.

We mention first a few technical terms. According to Hamilton, the quantity T is designated as the *tensor of the quaternion.* Thus one can say: *an ordinary rotation is a unit quaternion* (that is, a *quaternion with the tensor* 1).

We designate, further, the term D in the expression (6) as the *scalar part of the quaternion.* The term D is determined by a specified numerical quantity $\sqrt{T}\cos\frac{\omega}{2}$ that may be fixed through measurement on an arbitrary scale, and has no particular orientation in space. The term $iA + jB + kC$ is then called the *vector part of the quaternion.* Namely, this term has the character of a vector, since it has, at the same time, a specified axis (given through the angles a, b, c) and a specified magnitude (given through the product $\sqrt{T}\sin\frac{\omega}{2}$).

We consider, in particular, a quaternion that reduces to its vector

part. According to equations (1), this is a rotation-stretch with the rotation angle $\omega = \pi$; it consists of a stretch and a reversal of space around the axis a, b, c. We wish to use the word *turning-stretch*[38] for this operation. On the other hand, such a quaternion has the complete character of a vector;[39] its components A, B, C are directly the components of a line segment with length $\sqrt{T}$ on the axis a, b, c. We can thus say: *vectors appear in quaternion theory as turning-stretches.*

A vector (V) yields, for example, the binding of an arbitrary point (XYZ) with the origin O. Conceived as a turning-stretch, we can represent it as

$$V = iX + jY + kZ. \quad \text{—} \tag{7}$$

With the introduction of the quaternion notation, there now arises the further objective of making calculations with the symbolic expressions of the form (6) similar to calculations with the usual complex numbers. As a precondition for this, one must establish rules for calculations with quaternions. In themselves, such rules can naturally be assumed arbitrarily. To construct the quaternion calculus purposefully, however, the usual operations are adopted.

In this spirit, we first define the *addition of quaternions*, which, however, comes into consideration for us only incidentally. We allow ourselves to be led by the principle that the addition of quaternions should coincide, in the special case of turning-stretches, with the usual addition of vectors; that is, it should coincide with the parallelogram construction. This yields the following definition:

Two quaternions are added by adding the scalar parts on the one hand and the vector parts on the other hand; the former in the usual algebraic sense, and the latter according to the rule of geometric addition. If there is thus given, in addition to Q (see eqn. (6)), a second quaternion

$$Q' = iA' + jB' + kC' + D',$$

then their sum Q'' is the quaternion

$$Q'' = Q + Q' = i(A + A') + j(B + B') + k(C + C') + (D + D').$$

This procedure, moreover, can also be seen as the generalization of the addition of the usual two-term complex numbers.

We next define a procedure for *multiplication of quaternions.* The principle that guides us here should be the geometric representation of rotation-stretches. We establish:

Two quaternions are multiplied by composing the corresponding rotation-stretches according to the previous rule.

Thus if Q is a first and Q' a second quaternion, their product

$$Q'' = QQ'$$

should have as components exactly those quantities A'', B'', C'', D'' that are determined through equations (5). Since the components of Q and Q' in these equations appear in an asymmetric way, the result of the multiplication will be dependent on the order of the factors. Thus the theorem:

Multiplication of quaternions is not a commutative operation.

This theorem is, in our conception, an immediate consequence of the repeatedly emphasized fact that the resultant of multiple finite rotations depends on the order of the individual rotations.

We can give this multiplication rule yet another noteworthy expression. Namely, if we carry out the product QQ' purely formally, as if for the multiplication of ordinary multitermed algebraic expressions, then there appear terms with factors i^2, ij, $j^2, \dots$. If the resulting expression is to coincide with the product QQ' given above, then we must establish the meaning of each of these factors. *Let*

$$i^2 = j^2 = k^2 = -1, \tag{8}$$

$$ij = k, \quad jk = i, \quad ki = j, \tag{9}$$

and

$$ji = -k, \quad kj = -i, \quad ik = -j. \tag{10}$$

Then the formal calculation of the product QQ' reduces exactly to equations (5)*).

We remark, further, that equations (9) and (10), which, moreover, are not independent of each other, display again the noninterchange-

*) The composition formulas, and therefore the multiplication formulas (5), were, at least in the case $T = 1$, already known, in principle, before Hamilton; they are equivalent to the formulas that Rodrigues has given in Liouville's Journal (1. sér.) V, 1841 for the composition of rotations. To see this, one considers, instead of the quaternion components themselves, their ratios through the proportion

$$A : B : C : D = \lambda : \mu : \nu : 1.$$

The quantities λ, μ, ν already appear for the specification of an individual rotation in Euler; they will be designated (in contrast to φ, ψ, ϑ) as *Euler's symmetric rotation parameters*. If one now writes equations (5) in terms of λ, μ, ν, in that one divides the last of these equations into the preceding three, then the formulas of Rodrigues follow.

ability of the factors of a quaternion product. We can obviously consider the units i, j, k as special quaternions in themselves. Geometrically, they are turning-stretches of tensor 1; that is, reversals[40] about the x-, y-, z-axes, respectively. The formulas (8), (9), and (10) then easily follow geometrically, although at first only up to the sign. Namely, they simply state that the result of two reversals about the same axis is the identity; two reversals about two perpendicular axes give a reversal about the common perpendicular to the two axes. If one also wishes to obtain the signs correctly, then one must, according to page 36 and the following, go over from the consideration of the whole rotation angle to that of the half rotation angle. One then recognizes that $i^2 = -1$, since i^2 is a full rotation whose half-angle modulo 2π is equal to π (and not equal to zero). Moreover, formula (8) calls to mind the equation $i^2 = -1$ from the theory of ordinary complex numbers. —

We now enter in somewhat more detail into the multiplication of turning-stretches; that is, into the *vector product*, whose later use will, on occasion, serve us well. Let two vectors v and v' be given through the coordinates of their endpoints x, y, z and x', y', z', respectively. We then form

$$vv' = (ix + jy + zk)(ix' + jy' + kz').$$

The multiplication rules (8), (9), and (10) yield as the result a quaternion, namely

$$\begin{aligned} vv' = {} & i(yz' - zy') + j(zx' - xz') + k(xy' - yx') \\ & - (xx' + yy' + zz'). \end{aligned}$$

We consider the scalar and vector parts of this quaternion individually. We call the former, taken negatively, the *scalar product* of the vectors v and v' (in G r a s s m a n n's terminology, the "inner product"); the latter is called the *vector product* (equal to G r a s s m a n n's "exterior product").[41] The scalar product may be characterized by the symbol $S(v, v')$, and the vector product by the symbol $V(v, v')$.

The scalar and vector products may be understood so that they have a meaning independent of the choice of the coordinate system in which x, … and x', … are measured. This follows immediately from their introduction as the resultant of the two turning-stretches; we also recognize it, however, on the basis of their direct geometric meaning. *The scalar product, namely, is equal to the product of the length of v and the*

projection of v *into* v', as follows immediately from the formula

$$S(v, v') = xx' + yy' + zz'. \tag{11}$$

The vector product is itself a vector that is perpendicular to v *and* v' *in the direction from which* v *is seen to be transformed into* v' *by a clockwise rotation along the shortest path; the length of this vector is equal to the area of the parallelogram constructed from* v *and* v'. One also verifies this without difficulty on the basis of the formula

$$V(v, v') = i(yz' - zy') + j(zx' - xz') + k(xy' - yx'). \text{ —} \tag{12}$$

We now go further, and consider equations (2) in terms of the formal operation of quaternion multiplication. In equations (2), we can say, the two vectors

$$v = ix + jy + kz$$

and

$$V = iX + jY + kZ$$

that result from the respective joining of the points x, y, z and X, Y, Z with O are related to each other through the rotation-stretch

$$Q = iA + jB + kC + D.$$

We assume as the basis of the representation, in agreement with that of equations (2), that the direction of v lies, with respect to the direction V, in such a way that the former is transformed into the latter through the rotation contained in Q, and, further, that the length of V is related to the length of v so that the former is transformed into the latter through the stretch contained in Q.[42] If T is the tensor of Q and we take, as is convenient and amounts to no essential restriction, the length of V to be 1, then the length of v will be equal to T. In the following, we now conceive not only Q but also v and V as *operations*, the latter naturally as *turning-stretches*. We then arrive very elegantly at equations (2) in connection with the well-known representations of group theory in the following way.

We mark the axes of the turning-stretches v and V and the rotation-stretch Q extending from O. We represent the object on which the operations v, V, and Q act — in the same way as toward the end of §3 — by a unit sphere centered at O. In the initial position, the sphere may be denoted by K. On this sphere we conduct two series of operations, which will lead to the same result.

1) We subject K to the turning-stretch v, so that K is reversed about the axis of v and is enlarged in the ratio of $T:1$. We denote the resulting sphere of radius T by $(K)v$. We then apply the enlargement of the sphere in reverse, in that we apply to $(K)v$ a similarity transformation T^{-1} with enlargement proportionality $1:T$. Thus there results once again a unit sphere, which we denote by $(K)vT^{-1}$. We could also have obtained $(K)vT^{-1}$ from K, moreover, by a simple reversal about the direction of v.

2) Alternatively, we subject K successively to the rotation-stretch Q, the turning-stretch V, and the "inverse rotation-stretch Q^{-1}." By the latter words we mean that operation which makes the operation Q directly in reverse. This rotation-stretch Q^{-1} has the same axis as Q, produces the same rotation but in the opposite sense, and has the reciprocal stretch $\frac{1}{T}$. In particular, it transforms the direction of V into that of v.

In the rotation-stretch Q, K is first transformed into a sphere $(K)Q$ of radius T, where, in particular, those points that originally lie on the axis of v come to lie on the axis of V. We then apply the turning-stretch V. The sphere $(K)Q$ will thus be transformed into a sphere $(K)QV$ that is turned with respect to $(K)Q$ through two right angles and has the same radius as $(K)Q$. Finally, we add the rotation-stretch Q^{-1}. This operation may be applied, for the sake of clarity, to $(K)QV$ as well as to $(K)Q$. There result from $(K)QV$ and $(K)Q$ two unit spheres, which may be called $(K)QVQ^{-1}$ and $(K)QQ^{-1}$, respectively. $(K)QQ^{-1}$ is obviously identical to the initial sphere K. But we assert further that $(K)QVQ^{-1}$ coincides in magnitude and position with our earlier sphere $(K)vT^{-1}$. In fact, through the operation Q^{-1}, the mutual positioning of the spheres previously denoted by $(K)Q$ and $(K)QV$ is not changed. If $(K)Q$ is transformed into (K), then $(K)QV$ is transformed into a sphere that results from (K) by a reversal about that axis into which the axis of V is transformed by the operation Q^{-1}; that is, by a reversal about the axis of v.

But through this same reversal we have previously generated $(K)vT^{-1}$ from K. The operations described under 1) and 2) thus actually lead to the same result; that is, they give, taken together, the same rotation axis and the same rotation angle. The latter is at first determined only modulo 2π. To see that the rotation angle is also the same modulo 4π, we make a small continuity consideration. Namely, we let the angle $\frac{\omega}{2}$

corresponding to the quaternion Q increase from zero to its specified value. Our operations 1) and 2) certainly give the same rotation angle in the case $\frac{\omega}{2} = 0$. The equality will also occur throughout the steady increase to $\frac{\omega}{2}$. This continuity consideration is necessary because we would otherwise have undetermined signs in the resulting formula.

We express the equality of the operations 1) and 2) through the symbolic equation

$$(K)vT^{-1} = (K)QVQ^{-1},$$

or, if we remove the object K and regard only the operations,[43]

$$vT^{-1} = QVQ^{-1}. \text{ —} \tag{13}$$

We remark further that the two operations vT^{-1} and V have the same tensor (1) and the same rotation angle (π), and that our deliberation may be extended to any two rotation-stretches of coinciding tensor and rotation angle whose axes are transformed one into the other by means of Q. —

Equation (13) can now state, according to its derivation, nothing other than the equations (2); it must state the relation between the components x, y, z and X, Y, Z. In fact, one will return to the equations (2) if one takes the trouble to calculate the quaternion products correctly and equates the factors of i, j, and k on the right and left to each other individually.*). The calculation, however, is rather detailed. We thus wish to take the following simpler and, moreover, more instructive approach.

We add the operation Q on the right and left sides of equation (13), through which nothing is changed. We thus obtain

$$vT^{-1}Q = QV. \tag{14}$$

We now calculate, with consideration of (9), (10), and (11),

$$\begin{aligned} QV &= (iA + jB + kC + D)(iX + jY + kZ) \\ &= i(DX - CY + BZ) + j(CX + DY - AZ) + k(-BX + AY + DZ) \\ &\qquad - (AX + BY + CZ). \end{aligned}$$

*) This is remarked by C a y l e y, Philos. Magazine, Bd. 26, 1845; cf. Ges. W. Bd. 1, p. 123, where equation (13) appeared for the first time.[44]

In the same way there follows

$$\begin{aligned} vQ &= (ix + jy + kz)(iA + jB + kC + D) \\ &= i(Dx + Cy - Bz) + j(-Cx + Dy - Az) + k(Bx - Ay + Dz) \\ &\qquad - (Ax + By + Cz). \end{aligned}$$

Equation (14) thus separates into the four equations

$$(15) \qquad \begin{cases} \frac{1}{T}(\ \ Dx + Cy - Bz) = \ \ DX - CY + BZ, \\ \frac{1}{T}(- Cx + Dy + Az) = \ \ CX + DY - AZ, \\ \frac{1}{T}(\ \ Bx - Ay + Dz) = - BX + AY + DZ, \\ \frac{1}{T}(\ \ Ax + By + Cz) = \ \ AX + BY + CZ. \end{cases}$$

These equations have a remarkable structure; namely, in the first three equations the coefficients on the right and left sides each form a skew determinant, in which the horizontal rows of the one coincide with the vertical rows of the other.

It is now easy to attain, from equations (15), the equations contained in the schema (2). We solve equations (15) for x, for example, by multiplying the rows by D, $-C$, B, A and adding; the coefficient of x thus becomes

$$\frac{1}{T}(D^2 + C^2 + B^2 + A^2) = 1,$$

while the coefficients of y and z become zero; we finally obtain

$$x = (D^2 - C^2 - B^2 + A^2)X + 2(AB - CD)Y + 2(BD + AC)Z.$$

This, however, is one of the equations contained in (2). One finds the remaining equations in a corresponding manner. We can thus say:

The transformation equations (2) *are, from the standpoint of the quaternion theory, only a different way of writing the self-evident relation* (13).

As one sees, our geometric definition of rotation-stretches leads to a perfectly clear conceptual interpretation of the quaternion calculus. It has, moreover, the advantage of clearly delimiting the *range of application of quaternions.* As in the consideration above, quaternions will be suitable if one wishes to have a handy algorithm for the composition of rotations and stretches. We can summarize these operations as the

operations of the so-called "principal group" of geometry. They are characterized by the fact that they change the properties of geometric figures only in an inessential way; that is, only according to their magnitude and their scale. We will not, however, be able to attribute to the quaternion theory a more general significance for mathematics, in contrast to many advocates of the theory, who wish to view quaternions as an expedient for function theory and algebra that is just as fundamental as the use of the usual complex numbers. In general, the true *quaternion calculus* (in which additions and multiplications of quaternions are combined with each other in arbitrary multiplicity) appears to us to be of only secondary interest.

Moreover, the preceding remarks should not be interpreted to mean that quaternions are the most appropriate rotation parameters for our top problem, which does indeed treat essentially of the composition of rotations. In fact, the parameters α, β, γ, δ provide the advantage of greater analytic simplicity, and the Euler angles φ, ψ, ϑ provide greater geometric simplicity. If we will thus have, in the following chapters, no reason to return to the quaternion calculus, we will, in contrast, often make advantageous use of vector analysis. The addition of vectors has a natural place in mechanics on the basis of the theorem of the parallelogram of forces. We will also draw repeatedly upon the multiplication of vectors, since many formulas are expressed concisely through the language of vector analysis. —

As already mentioned at the beginning, our geometric definition of quaternions is not sufficiently emphasized in the usual presentations by H a m i l t o n and his successors. Hamilton, rather, defined a quaternion as "*the quotient of two vectors*." This definition is hardly appropriate as the basis of a theory; the expression "quotient of two vectors" first requires an explanation of its own, and refers us, in case such an explanation is not given, merely to an unclear analogy with the calculation rules of ordinary algebra. The terminology can naturally be justified theoretically, and may then even have a certain advantage to be mentioned below; but it appears inappropriate to begin with it.

In order to be able to attach a precise sense to Hamilton's definition, we will proceed in connection with equation (13) in the following way.

We first make clear that a rotation-stretch Q composed with a turning-stretch V, whose axis is perpendicular to the axis of Q, gives the same result as the succession of the turning-stretch V and the in-

verse rotation-stretch Q^{-1} combined with a similarity transformation of enlarging proportion T^2, where T is the tensor of Q. In symbols, this means

$$QV = VQ^{-1}T^2. \tag{16}$$

Here, as on page 62, V may, for the sake of brevity, denote a unit vector; that is, a simple reversal. It is possible to construct the proof of equation (16) geometrically by considering the successive transformations of a unit sphere K, as in the proof of equation (13). We wish to demonstrate the validity of equation (16) here on the basis of the quaternion multiplication laws. For this purpose, we place the x-axis in the direction of the axis of the turning-stretch V, so that V becomes simply equal to i. The first component of Q must then vanish, since the axis of Q is, by assumption, perpendicular to the axis of V. We thus have

$$V = i, \quad Q = jB + kC + D.$$

To find the inverse quaternion Q^{-1}, we change, according to page 63, the angle $\frac{\omega}{2}$ of Q to $-\frac{\omega}{2}$, the tensor T to $\frac{1}{T}$, and leave the axis unchanged. In consequence of these changes,

$$Q^{-1} = \frac{1}{T^2}(-jB - kC + D).$$

Equation (16) is thus transformed into the equation

$$(jB + kC + D)i = i(-jB - kC + D),$$

which, according to (8), (9), and (10), is identically satisfied.

We then substitute the value of VQ^{-1} from (16) into (13). There follows

$$Tv = Q^2V,$$

or

$$Q^2 = T\frac{v}{V}. \tag{17}$$

Since we assume that V is perpendicular to the axis of Q, the vector v that is obtained from V by the rotation must also be perpendicular to the axis of Q. We denote the product Tv by v'; then v' has length T^2, since, as we assumed, V has length 1. Further, we consider Q^2 as a new quaternion Q'; Q' then obviously has the same axis as Q, twice the rotation angle, and the square of the tensor of Q. Equation (17) is thus transformed into

$$Q' = \frac{v'}{V}. \tag{18}$$

Expressed in words, this equation yields Hamilton's definition of a quaternion. We can say:

The quaternion Q' *in equation* (18) *is represented as the quotient of the two vectors* v' *and* V *that are perpendicular to the axis of* Q'*, that form with one another an angle which is equal to half the rotation angle of* Q'*, and whose lengths are in proportion as the tensor of* Q' *to unity.*

This definition of a quaternion is obviously rather particular, and is a simplification of our original introduction of the concept. On the other hand, we dare not hide the fact that it has one great advantage over our definition. Namely, it shows immediately that it is important to give a unique specification to the *half* rotation angle $\frac{\omega}{2}$ of a quaternion, while our representation of the rotation-stretch operates first with the *entire* rotation angle ω, and only through the somewhat arbitrary stipulation of page 36 must the relation with the half rotation angle be established.

In conclusion, we wish to mention an advantage that applies, not without foundation, to quaternion theory and vector analysis; namely, the independence of their operations from the basis structure of the coordinate system.

The theory of quaternions, however, merely shares this advantage with any rational analytic geometry (whether it treats of metric relations with rectangular coordinates or of projective relations with homogeneous coordinates). What is always essential is that one *thinks invariantly*, not that one *calculates invariantly*.[45] It signified a misconception of the essence of analytic geometry when, as a matter of principle, one did not wish to use coordinates explicitly for the execution of calculations. In reality, quaternion theorists also use coordinates in every place where it appears useful to them; namely, each time they write the four components A, B, C, D of a quaternion explicitly.

Chapter II

Introduction to kinetics (statics and impulse theory)

§1. Contrast between continuously acting forces and impact forces; the impulse for a single free mass particle

While in the preceding kinematic chapter we had no occasion to speak of the principles of mechanics, we will have to draw upon these principles in some form for the kinetic considerations that now follow. This can be done in various ways.

The original setting of the principles of mechanics by Newton assumes the *concept of force* as something immediately acknowledged and understood. More recent presentations often seek to eliminate this concept from the foundations, and introduce it only later as a convenient abbreviated designation in mechanics. The suitability of one or the other procedure depends essentially on the objective that one pursues. If one sets out merely to construct a consistent conceptual system, as Hertz does in his beautiful work on the "Principles of Mechanics," then one can well dispense with the concept of force. If one wishes, however, to attain a lively comprehension of mechanical phenomena and a rapid orientation in specific questions, as is our intention in these lectures, then the concept of force appears particularly valuable on psychological grounds; namely, this concept is immediately associated with the activities of man, who has in his muscles the possibility of doing work. Such a performance of work is bound in our perception with the feeling of *exertion*. We involuntarily carry over this perception to the external predecessors of motion as well. The root of the concept of force lies, without doubt, in this anthropomorphic interpretation of external events. We will not, therefore, suppress this interpretation, but rather proceed with this point of view directly in the foreground.[46]

R.J. Nagem, G. Sandri, *The Theory of the Top*, DOI: 10.1007/978-0-8176-4721-6_II,

We determine whether a force acts in a certain direction on a mass particle, which is thought to be fixed somehow in space, by displacing the particle slightly in the opposite direction. If we must perform work (exert our muscles) in doing so, then a force is present; otherwise, the particle is free of force in the given direction. The same procedure can also serve for the measurement of force: *we measure a force acting on a particle in a specified direction by the ratio of the work that we must perform in a displacement of the particle in the opposite direction to this displacement* $\left(P = \frac{A}{l}\right)$. If the units of length and work are prescribed, then the unit of force is determined as well.

We have already defined *the unit of length* on page 11 when we accepted the "absolute system of measurement." *The unit of work* is defined in this measurement system, as is well known, by first establishing *the unit of mass* as the gram, and then referring to experience with freely falling bodies (the so-called law of falling bodies). The unit of work is determined in this way as the 980.60^{th} part of the work that is performed in the elevation of one gram by one centimeter at the 45^{th} degree of latitude.[47] Work thus receives the dimension $\frac{ml^2}{t^2}$, and force the dimension $\frac{ml}{t^2}$. We must certainly declare this definition of the unit of work to be rather indirect from the above psychological point of view. Since force and work lie more immediately in our perception than mass, we could demand that the unit of work be established before the unit of mass, and could advocate the introduction of a general measurement system (which, moreover, is suggested occasionally from other points of view) in which work is used in addition to space and time as the third fundamental dimension. Whether such a measurement system would be recommended practically, we leave here completely undecided.

One traditionally distinguishes, furthermore, two kinds of forces: *continuously acting forces* and *impact* or *instantaneous forces.* The rule given above for the measurement of forces refers, as we must explicitly emphasize, to continuously acting forces. To include the measurement of impact forces, we say that *an impact force is equivalent to an extraordinarily large continuous force that acts for only an extraordinarily short time.*

We then measure the magnitude of an impact force[48] by the *product of the magnitude of that continuously applied force and the length of the interval of application* $\left([P] = \dfrac{ml}{t}\right)$.

We now recall briefly the empirical facts (or axioms) that constitute the statics of forces (both continuous forces and impact forces) for a single mass particle. They can be summarized in one statement:

The most general system of forces on a single freely moving mass particle has the character of a vector emanating from the particle.

In this statement we comprise the theorem of the parallelogram of forces as well as the decomposition of a force into components; *forces add exactly as vectors.*

The vector representation of a force is so familiar to us that we can easily hold our axiom as self-evident and undervalue its significance. It is, therefore, perhaps not superfluous to consider the most immediate consequences of our axiom. A force has, just as a vector, a specified magnitude and a designated direction. We now introduce a mass particle that may be subjected to arbitrary external forces. The direction of the resulting force is opposite to the direction in which we must perform the maximum work for a displacement. If we displace the particle, on the other hand, in a direction deviating from this designated direction, then, according to our fundamental theorem, the work thus performed is equal to the previous maximum work multiplied by the cosine of the angle that the current displacement encloses with the previous. In particular, we will, whenever we must expend a certain amount of work in a specified displacement, acquire the same quantity of work in an oppositely directed displacement.

We next recall the axiom that lies at the foundation of the kinetics of a single mass particle. It is well known: *a continuously applied force causes an acceleration of the particle whose direction coincides with the direction of the force and whose magnitude, multiplied by the mass of the particle (using the previously accepted absolute system of measurement), is equal to the magnitude of the force;* further, *an instantaneous force causes an instantaneous change of velocity, whose direction coincides with the direction of the force and whose magnitude, multiplied by the mass of the particle, equals the magnitude of the impact force.*

Since the velocity and the acceleration of a particle have, just as the

force, the character of a vector, the equality of the vector components follows immediately from our axiom that states the equality of the vectors. If X and $[X]$ denote the x-component of a continuous and an impact force, respectively, then we have

$$(1) \qquad X = m\frac{d^2x}{dt^2} \text{ etc.,} \qquad\qquad (2) \qquad [X] = m\Delta\left(\frac{dx}{dt}\right) \text{ etc.}$$

Here $\Delta\left(\frac{dx}{dt}\right)$ simply denotes the change of $\left(\frac{dx}{dt}\right)$.

Moreover, impact forces can be attributed to continuously acting forces and conversely, as is demonstrated by well-known examples from physics. In the kinetic theory of gases, for example, the continuously applied pressure of a gas against the wall of a vessel is thought to be produced by the collisions of the gas molecules. The continuous pressure is thus resolved into a series of very small and very rapidly successive individual impacts. One proceeds conversely in the theory of elasticity if one wishes to follow, in detail, the impact of two spheres. The apparently instantaneous impact force is replaced by a very short but still continuously increasing and decreasing force that is applied from the impacting to the impacted sphere.

One recognizes immediately that equations (1) and (2) are transformed into each other by one or the other interpretation. If we imagine, as in the example from the theory of gases, a series of individual impacts $[X_1]$, $[X_2]$, …, $[X_n]$ that follow each other by the very short time interval Δt, and we assume that the ratio $[X_i]:\Delta t$ tends to a finite limit X with decreasing Δt, then we obtain equation (1) from equation (2) through the passage to the limit $\Delta t = 0$. If we begin, on the other hand, from a continuously applied force X that differs from zero only in the time interval Δt, but that has very significant intensity, so much so that the "time integral"

$$\int_{t_0}^{t_0+\Delta t} X\,dt$$

retains a finite value $[X]$ if we let Δt decrease to zero, then we obtain equation (2) from equation (1) by integration with respect to t.

Since impact forces and continuous forces are thus, in a certain sense, mathematically equivalent, it is possible to base mechanics

equally well on one or the other. Each of the two methods corresponds to a particular point of view of natural philosophy. Whoever is of the opinion that no discontinuous processes can occur in nature will prefer continuous forces. But whoever holds that continuity in nature is only apparent, and that our imperfect sensory organs yield an indistinct picture of the world, will generally wish to return to impact forces*). For the mathematician as such these questions do not come into consideration; the mathematician will evaluate the advantages of the two methods according to their greater or lesser mathematical usefulness and convenience.

From this point of view, we cannot refrain from giving preference to the method of impact forces over the now usual method of continuous forces for the general introduction of mechanics. We will also make use of this method when it is not at all a matter of treating sudden or rapidly successive changes in the state of motion. We thus take up again a manner of representation that was generally accepted by the actual founders of mechanics.

Fundamentally, moreover, it is only a matter of grasping the meaning of the differential equations of mechanics, or of differential equations in general, independently of formulas, according to their inner conceptual content.

We illustrate our intentions first with the example of a single freely moving mass particle.

We consider our particle at any position on its path, and always imagine *that impact force which is able to transform the particle instantaneously, and without change of position, from the state of rest into the actual state of motion.* If suddenly stopped at the considered place in its motion, the particle would be able to exert this same impact force against the obstacle. We call this impact force the *impulse of the particle*; our primary interest is now, to a certain extent, to follow not the motion of the particle, but rather the change in its impulse.

On the basis of the general axiom of statics we can say:

The impulse of a single free mass particle is a vector.

On the basis of the kinetic axiom we can, in addition, immediately give the magnitude and direction of this vector. According to equation (2), it is, namely, an impact force that changes the velocity 0 into the

*) Cf. here, for example, L. Boltzmann: Über die Unentbehrlichkeit der Atomistik in der Naturwissenschaft, Berichte der Wiener Akademie 1896, as well as Wied. Ann. 1897.[49]

velocity v, equal in magnitude to mv; the direction coincides with the direction of v. We will therefore say:

The magnitude of the impulse of a single mass particle equals the so-called quantity of motion;[50] *the direction of the impulse coincides with the actual direction of the advancement.*

Using the concept of the impulse, the motion of a single particle can now be described in a particularly simple way.

We have first

Theorem I. *If no forces act on the particle, the impulse remains constant in magnitude and direction in space.* (Galileo's law of inertia = Newton's lex prima).

If, on the contrary, the particle is subjected to external influences, these can consist of instantaneous impacts $[P_i]$ or a continuous force P. In the first case, the impacts $[P_i]$ are composed with the already present impulse according to the parallelogram of forces. In the second case, we construct the single impact force that is equivalent to the continuous force during the time element Δt; we thus form

$$[P] = \int_t^{t+\Delta t} P\,.\,dt = P\,.\,\Delta t.$$

This "infinitesimal impact" is again composed with the already present impulse according to the parallelogram of forces.

We thus formulate

Theorem II. *If external forces act on our particle, the impulse is altered so that its change at each moment Δt is equal in magnitude and direction to the (finite or infinitesimal) impact produced during this moment of time by the external forces* (Newton's lex secunda).

These fundamental theorems carry over word for word, as we will later see, to the case of the top, and mutatis mutandis to arbitrary mechanical systems. —

We next ask for *the work that the impulse or any continuous force must perform for the generation of the instantaneous state of motion; that is, the work that is necessary to transform the particle from the state of rest to the actually present state of motion.* This same work may be supplied by the particle in reverse if we arrest its motion suddenly or gradually. This explains the designation of our quantum of work as the *vis viva*[51] *of the particle.*

The work that a *continuous force* (X, Y, Z) performs on a mass particle in a displacement (dx, dy, dz) is evidently

$$dA = X\,dx + Y\,dy + Z\,dz = (Xx' + Yy' + Zz')\,dt. \tag{3}$$

This equation, in fact, is nothing more than the analytic expression for our original definition of the concept of force.

We substitute here for X, Y, Z the values from equation (1), in which we suppose that (X, Y, Z) is the force that has generated the actual motion of the particle beginning from rest. Then the element of work is

$$dA = m(x''x' + y''y' + z''z')\,dt.$$

The total work, which amounts exactly to the *vis viva* of the particle, follows through integration with respect to time; we thus obtain the well-known formula

$$T = \frac{m}{2}(x'^2 + y'^2 + z'^2). \tag{4}$$

The *vis viva* of the particle is thus determined by the momentary state of motion of the particle alone; it is independent of how we imagine this state to be produced, or in what way we let x', y', z' or X, Y, Z vary during the generation of the motion. It is therefore also valid, in particular, if we imagine the motion to be generated instantaneously by an impulse, as we now wish to do.

If we take up the calculation of the *vis viva* again under this special assumption, then the matter will become particularly clear. We can first suppose, for this purpose, that a force (X, Y, Z) with a constant and very large magnitude is applied during the very small time interval Δt, so that the velocity (x', y', z') increases *uniformly* during this interval. From our above expression for the work there then follows, if we denote the duration of the force by Δt,

$$T = \int_0^{\Delta t} dA = X\int_0^{\Delta t} x'\,dt + Y\int_0^{\Delta t} y'\,dt + Z\int_0^{\Delta t} z'\,dt.$$

At the beginning of the interval Δt, the velocity of the particle is equal to zero, and at the end it has become (x', y', z'). Since, moreover, it increases from its initial value to its final value uniformly, the time integrals on the right-hand side have the values $\frac{1}{2}x'\Delta t$, $\frac{1}{2}y'\Delta t$, $\frac{1}{2}z'\Delta t$. We thus have

$$T = \frac{1}{2}(Xx' + Yy' + Zz')\,\Delta t.$$

We now let Δt decrease to zero. Then the products $X\Delta t$, $Y\Delta t$, $Z\Delta t$ go over into the components of the impulse, which we denote by $[X]$, $[Y]$, $[Z]$:

$$[X]=\int_0^{\Delta t} X\,dt = X\Delta t,\ [Y]=\int_0^{\Delta t} Y\,dt = Y\Delta t,\ [Z]=\int_0^{\Delta t} Z\,dt = Z\Delta t.$$

As a result, we can write the latter expression for the *vis viva* as

$$T = \frac{1}{2}([X]x' + [Y]y' + [Z]z'). \tag{5}$$

The vis viva *generated by the impulse appears here as half the product of the magnitude of the impulse and the length of the velocity vector.*

The expression (5) is, naturally, identical with (4). In fact, it follows immediately from our definition of the impulse given above that

$$[X] = mx', \qquad [Y] = my', \qquad [Z] = mz'. \tag{6}$$

These equations can also be written in the noteworthy form

$$[X] = \frac{\partial T}{\partial x'}, \qquad [Y] = \frac{\partial T}{\partial y'}, \qquad [Z] = \frac{\partial T}{\partial z'}. \tag{7}$$

In the execution of the partial differentiation, T is to be in the form (4); that is, written as a function of the velocity components.

We can, on the other hand, also conceive T as a function of the components of the impulse. Namely, there follows from equation (4) and (6)

$$T = \frac{1}{2m}([X]^2 + [Y]^2 + [Z]^2). \tag{4$'$}$$

As a result, we can also give equations (7) the form

$$x' = \frac{\partial T}{\partial [X]}, \qquad y' = \frac{\partial T}{\partial [Y]}, \qquad z' = \frac{\partial T}{\partial [Z]}. \tag{7$'$}$$

In the formation of these equations, T is naturally defined by equations (4′); that is, T is conceived as a function of the impulse components. In so far as a misunderstanding is precluded by the context, as here, we will not specifically indicate by the notation whether we assume one or the other conception of the *vis viva*.

In addition to equations (7) and (7′), we present as analytic expressions for the force components X, Y, Z the equations

$$X = \frac{\partial A}{\partial x}, \qquad Y = \frac{\partial A}{\partial y}, \qquad Z = \frac{\partial A}{\partial z}, \tag{8}$$

which result, respectively, from our original introduction of the concept of force in equation (3)*).

It is an immediate consequence of our impulse theorems I) and II) that the *vis viva* remains constant ($dT = 0$) for force-free motion of the particle, and that the change in the *vis viva* for motion influenced by external forces is equal to the work performed by these forces ($dT = dA$).

Equations (7), or the equivalent equations (7′), give us relations between the velocity vector, the impulse vector, and the expression for the *vis viva*, which we bring together in words in the following way:

The impulse (velocity) components are the partial differential quotients of the vis viva *with respect to the velocity (impulse) components, where we think of the* vis viva *given as a function of the velocity (impulse) components.*

In addition to equations (7) or (7′), there is a second triplet of equations that indicate how the impulse is changed by external influences. This triplet of equations is only the analytic expression of the law stated in Theorem II. If a continuous force (X, Y, Z) acts on our particle, then we obviously have, according to II, with the use of rectangular coordinates,

$$\frac{d[X]}{dt} = X, \qquad \frac{d[Y]}{dt} = Y, \qquad \frac{d[Z]}{dt} = Z. \tag{9}$$

Formulas (7), (8), *and* (9) *are the very well known fundamental equations of particle mechanics in rectangular coordinates.* —

We now wish to ask how these equations change if we introduce, instead of *rectangular coordinates*, an arbitrary set of *generalized coordinates*. There is, to be sure, no reason for departing from rectangular coordinates for a single freely moving mass particle. The following considerations, however, should serve as a preparation for more difficult cases in which we cannot manage with rectangular coordinates.

We wish to think of the position of a particle in space as given not by three mutually perpendicular planes, but rather by three arbitrary surfaces, and to consider as coordinates not the quantities x, y, z, but rather any three functions $\varphi = \varphi(x, y, z)$, $\psi = \psi(x, y, z)$, $\vartheta = \vartheta(x, y, z)$. Instead of the usual velocity coordinates x', y', z', we introduce the

*) It is noted that the meaning of the differential symbols in (8) deviates from the usual meaning, in so far as they are indeed "differential quotients," but are not to be used as "derivatives" of a function of the coordinates, since the expression (3) does not, in general, represent a perfect differential.

corresponding "generalized velocity coordinates" φ', ψ', ϑ'; that is, the derivatives of our quantities φ, ψ, ϑ with respect to time.

One recognizes immediately, if one differentiates the defining equations for φ, ψ, ϑ with respect to time, that *the new velocity coordinates are linear functions of the old, and vice versa.* In particular, we wish to denote the coefficients of φ', ψ', ϑ' in the linear expressions for x', y', z' by a_{ik}, so that

$$(10) \qquad \begin{cases} x' = a_{11}\varphi' + a_{12}\psi' + a_{13}\vartheta', \\ y' = a_{21}\varphi' + a_{22}\psi' + a_{23}\vartheta', \\ z' = a_{31}\varphi' + a_{32}\psi' + a_{33}\vartheta'; \end{cases}$$

the meaning of the coefficients a_{ik} is evidently

$$(11) \qquad a_{11} = \frac{\partial x}{\partial \varphi}, \quad a_{12} = \frac{\partial x}{\partial \psi}, \cdots \; a_{21} = \frac{\partial y}{\partial \varphi}, \; \cdots .$$

It will now be further necessary, however, to consider also *the corresponding generalized coordinates of the force and the impulse.* To define these coordinates, we return to our original definition of force. We ask for the work that we must perform in an infinitesimal change of the coordinate φ, with fixed values of ψ and ϑ. We define the φ coordinate of the force as the ratio of this work to the change in φ. If we denote this coordinate of force by Φ, then we have

$$\Phi = \frac{\partial A}{\partial \varphi}$$

(where the meaning of this differentiation symbol is defined by the conditions $\psi = \text{const.}$, $\vartheta = \text{const.}$); the force components Ψ and Θ have corresponding meanings. The expression for the work in an arbitrarily small displacement $(d\varphi, d\psi, d\vartheta)$ of our mass particle thus becomes

$$(12) \qquad dA = \Phi\, d\varphi + \Psi\, d\psi + \Theta\, d\vartheta = (\Phi\varphi' + \Psi\psi' + \Theta\vartheta')\, dt.$$

Our definition of the generalized force components thus implies that the expression for the work retains exactly the earlier form (3) *with the introduction of generalized coordinates.*

We can, therefore, easily express $\Phi\,\Psi\,\Theta$ in terms of XYZ. Namely, if we replace x', y', z' in (3) by their values in terms of φ', ψ', ϑ' given in equations (10) and order according to the latter quantities, then Φ, Ψ, Θ become the coefficients of φ', ψ', ϑ', respectively. We thus obtain

$$(13)\qquad \begin{cases} \Phi = a_{11}X + a_{21}Y + a_{31}Z, \\ \Psi = a_{12}X + a_{22}Y + a_{32}Z, \\ \Theta = a_{13}X + a_{23}Y + a_{33}Z. \end{cases}$$

The expressions for the new force coordinates in terms of the old are thus entirely analogous to the expressions for the old velocity coordinates in terms of the new; namely, the coefficients of the force transformation result from the coefficients of the velocity transformation by the interchange of the horizontal and vertical ranks. To state this concisely, we say:

The force coordinates are contragredient to the velocity coordinates.

The coordinates of an impulse, which we can indeed conceive as the limiting case of a continuous force, behave just like the coordinates of a continuous force; further, the expression for the *vis viva*, which we have defined as a certain finite quantity of work, transforms just like the expression for the infinitesimal work.

We thus obtain

$$(14)\qquad \begin{cases} [\Phi] = a_{11}[X] + a_{21}[Y] + a_{31}[Z], \\ [\Psi] = a_{12}[X] + a_{22}[Y] + a_{32}[Z], \\ [\Theta] = a_{13}[X] + a_{23}[Y] + a_{33}[Z] \end{cases}$$

and

$$(15)\qquad T = \frac{1}{2}\left([\Phi]\varphi' + [\Psi]\psi' + [\Theta]\vartheta'\right).$$

On the basis of equations (14), we verify without difficulty the relations

$$(16)\qquad [\Phi] = \frac{\partial T}{\partial \varphi'}, \qquad [\Psi] = \frac{\partial T}{\partial \psi'}, \qquad [\Theta] = \frac{\partial T}{\partial \vartheta'},$$

in which we suppose T to be expressed as a function of the velocity coordinates. In fact, it follows from $T = \frac{m}{2}(x'^2 + y'^2 + z'^2)$, for example, that

$$\frac{\partial T}{\partial \varphi'} = mx'\frac{\partial x'}{\partial \varphi'} + my'\frac{\partial y'}{\partial \varphi'} + mz'\frac{\partial z'}{\partial \varphi'} = [X]a_{11} + [Y]a_{21} + [Z]a_{31} = [\Phi].$$

Equations (16) are precisely analogous to our earlier relations (7). *The relations* (7) *are completely unchanged in form by the introduction of generalized coordinates.*

The same holds also for equations (7′); we assure ourselves of this briefly in the following way.

Solved for $\varphi', \ldots$, equations (10) give

$$(10') \qquad \begin{cases} \varphi' = A_{11}x' + A_{21}y' + A_{31}z', \\ \psi' = A_{12}x' + A_{22}y' + A_{32}z', \\ \vartheta' = A_{13}x' + A_{23}y' + A_{33}z', \end{cases}$$

where the A_{ik} generally denote the subdeterminants of the determinant a_{ik} divided by this determinant. In the same way, there result from equations (14)

$$(14') \qquad \begin{cases} [X] = A_{11}[\Phi] + A_{12}[\Psi] + A_{13}[\Theta], \\ [Y] = A_{21}[\Phi] + A_{22}[\Psi] + A_{23}[\Theta], \\ [Z] = A_{31}[\Phi] + A_{32}[\Psi] + A_{33}[\Theta]. \end{cases}$$

We now imagine that T is expressed in terms of $[\Phi]$, $[\Psi]$, $[\Theta]$ by substituting the values of $[X]$, $[Y]$, $[Z]$ from (14′) into equation (4′). We then form $\frac{\partial T}{\partial[\Phi]}$, in which we hold $[\Psi]$ and $[\Theta]$ as well as φ, ψ, ϑ fixed. We thus have, with consideration of (7′) and (14′),

$$\frac{\partial T}{\partial[\Phi]} = \frac{\partial T}{\partial[X]} \cdot \frac{\partial[X]}{\partial[\Phi]} + \frac{\partial T}{\partial[Y]} \cdot \frac{\partial[Y]}{\partial[\Phi]} + \frac{\partial T}{\partial[Z]} \cdot \frac{\partial[Z]}{\partial[\Phi]} = A_{11}x' + A_{21}y' + A_{31}z'.$$

But from this it follows, according to (10′), if we adjoin at the same time the analogous equations,

$$(16') \qquad \varphi' = \frac{\partial T}{\partial[\Phi]}, \quad \psi' = \frac{\partial T}{\partial[\Psi]}, \quad \vartheta' = \frac{\partial T}{\partial[\Theta]}.$$

If we wish to formulate equations (16) and (16′) as theorems, then the manner of expression of page 77 can serve us word for word.

We then have in analogy to equations (8), according to (12),

$$(17) \qquad \Phi = \frac{\partial A}{\partial \varphi}, \quad \Psi = \frac{\partial A}{\partial \psi}, \quad \Theta = \frac{\partial A}{\partial \vartheta}.$$

For the meaning of these differential symbols, cf. the footnote on p. 77.

It will also be well to rewrite equations (9) in the generalized coordinates φ, ψ, ϑ. For this purpose, we first multiply equations (9) sequentially by a_{11}, a_{21}, a_{31} and add them. Then there results on the right-hand side, according to (13), the component Φ of the external force. We write the left side as

$$\frac{d}{dt}(a_{11}[X] + a_{21}[Y] + a_{31}[Z]) - \left([X]\frac{da_{11}}{dt} + [Y]\frac{da_{21}}{dt} + [Z]\frac{da_{31}}{dt}\right).$$

The first term here is simply the differential quotient of the impulse component $[\Phi]$ with respect to time; the second term becomes equal, with consideration of (6) and (11), to

$$m\left(x'\frac{\partial x'}{\partial \varphi} + y'\frac{\partial y'}{\partial \varphi} + z'\frac{\partial z'}{\partial \varphi}\right);$$

this is, however, nothing other than the partial differential quotient of the *vis viva* taken with respect to φ, where we think of the *vis viva* expressed in terms of the velocity coordinates φ', ψ', ϑ'. We thus arrive at the following law for the change of the φ-component of the impulse:

$$\frac{d[\Phi]}{dt} - \frac{\partial T}{\partial \varphi} = \Phi. \tag{18}$$

In precisely the same way there result

$$\begin{cases} \dfrac{d[\Psi]}{dt} - \dfrac{\partial T}{\partial \psi} = \Psi, \\[2mm] \dfrac{d[\Theta]}{dt} - \dfrac{\partial T}{\partial \vartheta} = \Theta. \end{cases} \tag{18}$$

Equations (18) state nothing other, according to their derivation, than equations (9); the simplicity of the general law II is only somewhat veiled here by the introduction of the coordinates φ, ψ, ϑ.

Equations (7), (8), and (9) or equations (16), (17), and (18) represent, taken together, the equations of motion for a single mass particle. One has, in the notation we have chosen, the simplest case of the so-called *Lagrange equations of the second kind.*[52] We will often return to these equations, and already remark here that we will always succeed, with help of an impulse concept analogous to the above, in interpreting them in a manner similar to the equations of motion of a single mass particle.

We have taken the term "impulse" from the work of Thomson and Tait, in which our concept plays an important role. Maxwell applies the same term in the attempt to use energy as the basis of the general equations of mechanics. The somewhat colorless word *momentum* is usually used in English books instead of impulse;[53] the components of the impulse are then called "the moments of momentum"(!). Hertz, on the other hand, uses the word *moment* as synonymous with our impulse. The otherwise very common designation "quantity of motion" (quantité de mouvement) expresses only the length but not the direction of the impulse vector, and therefore appears to us inappropriate.

§2. The elementary statics of rigid bodies

Before we can attack the kinetics of the top, we must orient ourselves with respect to the composition and decomposition of a system of forces applied to our body. One groups, as is well known, all those

investigations that treat merely of forces, without regard to the resulting motion, under the rather inappropriate name of *statics*, in which the question of the composition of a given system of forces is reduced to the search for such forces which, added to the given system of forces, would produce equilibrium. The word *dynamics* would be more suitable; that word is usually applied, however, to the part of mechanics that we designate as kinetics.

The treatment of statics for the present case of the rigid body can be accomplished according to two essentially different methods, which are characterized by the names of Poinsot and Lagrange. We wish to refer briefly to both methods.

The statics of rigid bodies in the geometric treatment of Poinsot is based on a series of *axioms* that we recognize, in part, from the previous case of a single mass particle. We said that the force on a single mass particle has the character of a vector, and that multiple forces applied at the same point add as vectors. For rigid bodies the following axiom is added: *the point of application of a force can be displaced arbitrarily in the direction of the force.* This axiom is obviously independent of the previous, since its validity is bound essentially to the nature of a rigid body; it can be regarded directly as the definition of the latter.[54] For actual bodies, which are always elastic to a certain degree, the axiom is naturally fulfilled only approximately. Moreover, Newton's lex tertia, which states the equality of actio and reactio, is interpreted so that it comprises our axiom, which appears somewhat artificial indeed.

With the help of this axiom, one now investigates the composition of forces which, somehow given, are distributed on the body. One first sees immediately that two forces that act in parallel directions can always be replaced by a single force whose direction is parallel to the directions of the original forces. The determination of the point of application and the magnitude of this single force forms the content of the so-called "law of the lever." If the forces are equal in magnitude and opposite in direction, a remarkable singularity results. Namely, the point of application then moves to infinity, while, at the same time, the magnitude become infinitely small.[55] *A force-pair* (that is, a pair of

oppositely directed equal and parallel forces) *is thus equivalent to an infinitely small force that acts on an infinitely long lever arm.*

To keep the presentation elementary, however, the consideration of infinitely small forces and infinitely large lever arms is usually avoided. As a result, one is obliged to regard the force-pair as an irreducible element of the statics of rigid bodies. It becomes necessary, further, to discuss the composition and decomposition of force-pairs in a manner similar to the composition and decomposition of forces. New axioms are not needed for this purpose, since the very definition of the force-pair allows the question of the equilibrium of pairs to be reduced to the question of the equilibrium of forces.

We summarize the results of these investigations in the following way: we represent the pair by a *vector*, which we lay down perpendicular to the plane lying through the forces, on that side from which the forces appear to act in the clockwise sense. The length of the vector (in the centimeter scale chosen once and for all) is equal to the "moment of the pair"; that is, equal to the product of the magnitude of the forces and their shortest distance of separation. The initial point of our vector can be taken arbitrarily in the plane of the pair, or also arbitrarily in space. Thus obtains the theorem: *two force-pairs are composed in such a way that the corresponding vectors add geometrically. Composition of multiple pairs always results again in a pair.*

Force-pairs thus have, just as forces, the character of vectors. We must, however, emphasize the following difference. The vector of a force applied to a rigid body may be displaced only in its direction, while the vector of a pair, in contrast, may be transported parallel to itself arbitrarily in space. *The vector of a force is* (in the manner of expression of Mr. B u d d e) *a sliding*[56] *vector, while that of a pair is a free vector; that is, a vector with a completely arbitrary point of attachment.*

It is evident from the preceding that a specified force-pair can be replaced, with regard to its static effect, in a great variety of ways by another force-pair. In fact, two force-pairs that give the same vector by the given construction are completely equivalent. It is thus advisable to abandon the particular manifestation of the force-pair and retain only the representing vector. We wish to accommodate this circumstance in the designation as well, and prefer to speak of a *turning-force* (or a

turning-moment) instead of a force-pair. We also designate the direction of the turning-force, according to its definition as a vector, as the "axis of the turning-force"; we mark the sense of the vector with an arrow that surrounds the axis in the sense that the forces of the pair appear to act as seen from the axis. The dimension of the turning-force is force times lever arm $\left(D = m\frac{l^2}{t^2}\right)$. In contrast to the expression turning-force (force-pair), we will temporarily use the word *pushing-force* (single-force).

We must beware, however, of associating the expression turning-force with the perception that the turning-force would tend to turn about a determined straight line. The turning-force occurs here in a purely static way. Its kinetic effect can be discussed only later, when we have treated of the kinetic effects of forces in general, and have made certain assumptions about the mass distribution of the body. The designation pushing-force should, naturally, imply equally little that the kinetic effect of a pushing-force necessarily consists in a parallel displacement.

We now enter into the general problem of statics, and therefore assume that forces are spatially distributed on the rigid body in an arbitrary manner. We proceed, as usual, in the following way: we adopt an arbitrary point O as the *reference point*, and place vectors through this point with the same and opposite senses as each of the given vectors. We group each of the given forces with the oppositely directed force through O into a force-pair, and replace the force-pair, according to the rule above, with the vector of a turning-force, where we conveniently choose the reference point O as the initial point of the vector. We thus obtain as many turning-forces as the number of original forces. We compose all these turning-forces into a resultant turning-force D, whose axis may pass though O. There then remain the forces (pushing-forces) through O in the same direction as the given forces. We compose these also into a resultant S. Thus the theorem:

An arbitrary system of forces applied to our rigid body may be replaced by a combination of a pushing-force and a turning-force emanating from an arbitrary point O.

We note the general rule for the calculation of S and D. Let P_i be one of the forces applied to our body, and P_i^x, P_i^y, P_i^z the projections

of the vector P_i on the axes of a rectangular coordinate system whose origin coincides with the reference point. Further, let X_i, Y_i, Z_i be the coordinates of the application point of P_i, and S^x, S^y, S^z and D^x, D^y, D^z the components of S and D, respectively. We then have

$$(1)\quad \begin{cases} S^x = \Sigma\, P_i^x, & S^y = \Sigma\, P_i^y, \qquad\qquad S^z = \Sigma\, P_i^z, \\ D^x = \Sigma\,(P_i^z Y_i - P_i^y Z_i), & D^y = \Sigma\,(P_i^x Z_i - P_i^z X_i), \ D^z = \Sigma\,(P_i^y X_i - P_i^x Y_i). \end{cases}$$

These very well known formulas are the immediate analytic expression of the geometric construction for S and D described above. One concisely calls S^x, S^y, S^z, D^x, D^y, D^z *the coordinates of the force system* (applied to the rigid body).

In general, the direction of S and the axis of D will form an angle that depends on the choice of the reference point and the nature of the given force system. But we can always choose the point O so that the vectors S and D coincide exactly in their directions, and O can still be chosen arbitrarily on a certain line. We will call this simplest equivalent of a general system of forces—a pushing-force combined with a turning-force that has the direction of the pushing-force as its axis—a *screw* (or, more precisely, a *force-screw*). We can then designate the quantities $S^x, \ldots, D^x, \ldots$ as *the coordinates of the force-screw* and can make the preceding theorem more precise in the following way:

An arbitrary force system applied to our rigid body may always be conceived as a screw whose coordinates are determined by (1).

If one places the reference point on the axis of the screw, then the components of S become proportional to those of D. However, one will often forgo this simplification in the formulas, and will prefer to give the reference point a position distinguished by the nature of the problem; this will, naturally, not hinder us from still representing the force system as a screw (although a screw not passing through O). Thus one prefers to choose the reference point for the freely moving rigid body as the center of gravity; we will later make this choice when we treat of the top moving on a plane. On the other hand, it is generally advisable to place the reference point at the fixed support point for the case of the (generalized or symmetric) top. If we construct the turning-force

and the pushing-force at this point O, we will then have, in the following, to consider only the turning-force.

In fact, whatever the details of the circumstances that effect the fixed position of the support may be, they must in any case provide, in opposition to the applied pushing-force at O, an equally large oppositely directed resistance force. This resistance force is called the *reaction force of the support point.* It is equal in magnitude to the resultant S found above, and is opposite in direction. If we add this reaction force to our system of forces, then the pushing-force S is directly canceled, while the turning-force remains exactly the previous quantity D. We can thus completely disregard the appearance the pushing-force from the start, and will have to return to it later only occasionally, if we would calculate the force that the foundation of the top must bear due to its motion.

At the same time, our general theorems above simplify in the present case. We can say:

The most general system of forces applied to our top can be replaced, with consideration of the fixed position of the support point, by a single turning-force. —

One observes the beautiful analogy between our static theorems and the kinematic theorems developed at the beginning of these lectures. The analogy for a free body is such that turning-forces must be compared with (infinitesimal) parallel displacements, and pushing-forces with (infinitesimal) rotations. The same geometric figure, the screw, appears once as a motion-screw, and once again as a force-screw. For the top, turning forces and (infinitesimal) rotations about O appear directly in parallel. Both are represented by vectors.

It is remarked, further, that the top is not inferior in simplicity, with respect to statics, to a single mass particle. The possibility of the later elementary geometric development for the theory of the top rests essentially on this circumstance.

As an example, we discuss the particularly simple case in which the original force system is provided by gravity. On each element dm of the top, the gravitational force $g\,dm$ acts vertically downward.

We take an X, Y, Z coordinate system with the origin at O and the Z-axis vertically upward. From equations (1) there follow immediately, if m is the total mass of the top and ξ, η, ζ are the coordinates of the center of gravity,

$$S^x = 0, \quad S^y = 0, \quad S^z = -g\int dm = -mg,$$

$$D^x = -g\int Y\,dm = -mg\eta, \quad D^y = g\int X\,dm = mg\xi, \quad D^z = 0.$$

We thus arrive at the very well known fact that *the effect of gravity is the same as a single force of magnitude mg applied perpendicularly downward at the center of gravity.*

For the symmetric top, the center of gravity obviously lies either on the figure axis or on its extension through point O. Let E be the

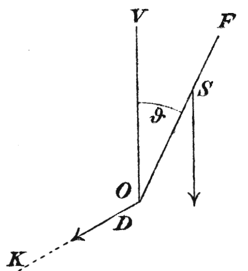


Fig. 11.

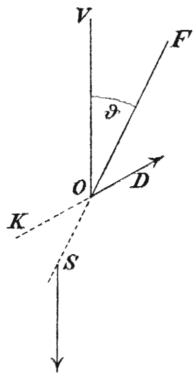


Fig. 12.

distance of the center of gravity from O, and ϑ, as earlier, the angle between the figure axis and the vertical. Then the vector D has length

$$mgE\sin\vartheta;$$

its direction is perpendicular to the vertical as well as to the figure axis. If we recall, further, the definition of the line of nodes on page 17, then we can say that the vector D lies on the line of nodes or on its extension through O, according to whether S lies on the figure axis or on its extension. We can also express this by saying that the vector D always lies on the line of nodes, and has the magnitude

$$D = P\sin\vartheta, \quad P = \pm mgE,$$

where the upper or lower sign is chosen according to whether the center of gravity lies above or below the support point (sc. for a vertically erected figure axis). The latter manner of expression, which we will later accept, has the advantage that we first treat of the two different

cases uniformly, and can separate them from one another afterward through the simple conditions $P > 0$ and $P < 0$.

As was the case in the latter example, it is usually assumed implicitly that the forces which are treated in statics are continuously acting forces. But one sees immediately that our entire development remains valid for impact forces, in so far as the collected impacts are applied only at the same time. In fact, each statement about continuous forces carries over immediately to impact forces (cf. the definition of the latter on page 70). In analogy to the concept of a turning-force, we will introduce the concept of a *turning-impact*; that is, the embodiment of a pair of equally large, oppositely directed parallel *pushing-impacts*. The magnitude, axis, and sense of the turning-impact are determined, just as for the turning-force, by the moment and the position of the constitutive single forces. The dimension of the turning-impact is $[D] = m\frac{l^2}{t}$. We therefore state the general theorems:

The most general system of impact forces applied to a freely moving rigid body may always be replaced by a single screw (more precisely said, an impact-screw).

and

The most general system of impact forces that is applied in any way to the points of our top may always be conceived as a single turning-impact, and may therefore be represented by a single vector emanating from O.

As mentioned previously, statics was founded in the form considered thus far by Poinsot (in whose work, however, the word "screw" does not appear). His fundamental work *éléments de statique* first appeared in the year 1803; an extraordinarily large number of editions has since followed. The proofs of the previously given theorems can be reviewed there. Statics was placed in relation with projective geometry, and especially with the linear geometry of Plücker, by Moebius.[57] In addition to the previously referenced theory of screws by Ball, we cite in particular, among more recent presentations, the two-volume textbook of Routh*), *Analytical Statics*, which may be especially recommended for precision and richness of content. —

We now enter into the previously mentioned second method of establishing the fundamental principles of statics, which is, in essence, due to Lagrange. It has, in contrast to the previously discussed

*) Cambridge, 2nd edition 1896.

presentation of Poinsot, the advantage of greater capacity for generalization; it may, on the other hand, appear less elementary. In this method, we derive the *composition of forces* from the *composition of the quantities of work* that the forces perform in an infinitesimal displacement of the rigid body.

We imagine again that an arbitrary system of forces P_i with arbitrary application points acts on our body. We compose the total work dA of our force system from the work elements dA_i of the individual forces P_i. We use the general fundamental theorem that *the work is independent of the choice of coordinate system; it is a scalar quantity; furthermore, quantities of work compose, as do scalar quantities, by addition in the algebraic sense.*

According to page 75, the work that the force P_i performs in an infinitesimal displacement of its point of application is

$$dA_i = (P_i^x x_i' + P_i^y y_i' + P_i^z z_i')\,dt,$$

where P_i^x, P_i^y, P_i^z, are the components of P_i and x_i, y_i, z_i are the coordinates of its point of application with respect to a coordinate system (x, y, z) fixed in space. Thus the total work is

$$dA = \Sigma\, dA_i = \Sigma\,(P_i^x x_i' + P_i^y y_i' + P_i^z z_i')\,dt.$$

We now recall the results of the first chapter, according to which every infinitesimal change of position of a rigid body consists of a parallel displacement and a rotation, and can be represented analytically by the six velocity coordinates x', y', z', p, q, r (cf. page 47). Using these coordinates, we first express the velocity (x_i', y_i', z_i') of the application point of P_i. As a result of the parallel displacement, the application point of P_i receives (just as every point of the body) the velocity (x', y', z'); as a result of the rotation, it acquires (see eqn. (3′) of page 41) the velocity

$$(-Y_i r + Z_i q,\ -Z_i p + X_i r,\ -X_i q + Y_i p).$$

Thus the resultant velocity of the application point of P_i is

$$\begin{aligned} x_i' &= x' - Y_i r\ + Z_i q, \\ y_i' &= y' - Z_i p + X_i r, \\ z_i' &= z' - X_i q + Y_i p. \end{aligned}$$

We enter these values into our above expression for the total work. This work is then written as

$$dA = (S^x x' + S^y y' + S^z z' + D^x p + D^y q + D^z r)\, dt, \tag{2}$$

where

$$\left\{\begin{aligned} &S^x = \Sigma\, P_i^x, \quad S^y = \Sigma\, P_i^y, \quad S^z = \Sigma\, P_i^z, \\ &D^x = \Sigma\,(P_i^z Y_i - P_i^y Z_i), \quad D^y = \Sigma\,(P_i^x Z_i - P_i^z X_i), \\ &D^z = \Sigma\,(P_i^y X_i - P_i^x Y_i). \end{aligned}\right. \tag{3}$$

These are precisely the quantities that appeared above in equations (1).

We wish to clarify the meaning of these quantities independently of what has been said earlier, and thus arrive at a simple new definition of them. According to equation (2), S^x is the ratio of the work that our system performs in a displacement of the body in the direction of the x-axis to the magnitude of this displacement. In the same way, D^x equals the ratio of the work that our system of forces performs in a pure rotation of the body about the X-axis, or, which is the same, the x-axis, to the magnitude of the rotation angle. The quantities S^x, $\ldots$, D^x, $\ldots$ therefore have meanings that are completely analogous to the meanings of the components P^x, $\ldots$ of the force P applied to a single mass particle, which indeed, for their part, were originally defined (cf. page 70) as the ratio of a certain quantity of work to a certain infinitesimal motion. As a result, the extension of the concept of force from the single mass particle to our rigid body is immediate. We are able to speak briefly of a *total force applied to the rigid body, which is equivalent to the given system of individual forces. This total force is decomposed into a pushing-force S and a turning-force D, each of which can be resolved into three components with respect to the coordinate axes.* We will again designate the quantities S^x, S^y, S^z, D^x, D^y, D^z as the coordinates of our total force, as we have named the quantities x', y', z', p, q, r the coordinates of the instantaneous velocity. We can then say in brief:

The coordinates of the force are, according to their definition, nothing other than the factors that multiply the coordinates of the velocity in the expression for the work.

If, in particular, we treat of a body with a fixed support point, which we will take, as previously, to be the reference point, then we have for the pushing velocity $x' = y' = z' = 0$, and we can abstract the pushing-force S^x, S^y, S^z. The turning-force, in contrast, will

again be defined precisely by equation (3). The total work that this turning-force performs in the infinitesimal displacement $(p, q, r)\,dt$ is therefore

$$dA = (D^x p + D^y q + D^z r)\,dt. \tag{2'}$$

The present definitions of the turning and pushing-force are, in many respects, preferable to the earlier ones, since they relate immediately to the concept of force for a single mass particle; they relieve us, in particular, from the introduction of the concept of the force-pair, which we earlier encountered by necessity. That the new and old definitions come to the same result is shown by the comparison of the expressions (1) and (3).

We could now develop anew the collected lessons of elementary statics; in particular, the facts that pushing-forces and turning-forces are sliding and free vectors, respectively, would appear as immediate consequences of the corresponding composition law of velocities and our fundamental theorem above, according to which quantities of work add as scalars.

Statics was established in this manner by Lagrange in his famous *mécanique analytique*. If we place the expression for work before the composition of forces, this is the same, in essence, as if we evaluate the equilibrium of a system on the basis of the *principle of virtual displacements*, according to the precedent of Lagrange. In fact, this principle states, as is well known, that an arbitrary system of given forces is in equilibrium if the work performed by the system vanishes for any possible infinitesimal displacement, or, somewhat more generally stated, that two different force systems are equivalent if the work performed by each of the two systems is equal for any possible displacement. In conformity with this principle, we have replaced the given force system of the P_i by a combination of a pushing-force S and a turning-force D. The means of expression was merely somewhat different than for Lagrange, in whose time the concept and the designation of work were not yet prevalent.[58] —

We wish, finally, to use the expression for the work to establish a convention for what we mean by the "*generalized coordinates of a force system*," just as was done for a single mass particle. We will proceed in complete analogy to page 78 for the single mass particle.

We can obviously establish the state of motion of the rigid body by many sets of six parameters other than the quantities x', y', z', p, q, r.

The nearest modification would be to select a different reference point, and to furthermore vary the positions of the xyz system in space and the XYZ system in the body. If we then decompose a specific force system into a pushing-force and a turning-force according to the above rule, we will find values for the components of these forces that differ from the previous. The coordinates of an infinitesimal change of position will change just as well. It is clear, on the other hand, that the work which a specific force system performs in a specific infinitesimal change of position must retain exactly the previous value. For a fixed choice of the units of length, time, and mass, the work has a fixed numerical value that is independent of the choice of coordinate system; it is (with respect to a change of the coordinate system) an *absolute invariant*, as we can say.

But we will consider still further changes in the coordinate specification of the instantaneous state of motion. We will establish, for example, the instantaneous rotation by the rates of change φ', ψ', ϑ' of the Euler angles instead of the quantities p, q, r, and we can further specify (as, for example, in the previous section) the position and velocity of the reference point in terms of the magnitudes and magnitude changes of three curvilinear coordinates ξ, η, ζ. The most general assumption is that we set x', y', z', p, q, r equal to arbitrary linear functions of six arbitrary velocity parameters ξ', η', ζ', φ', ψ', ϑ', with coefficients that depend on the values of ξ, η, ζ, φ, ψ, ϑ themselves. It is then asked how the coordinates of the force system are changed, or, more correctly, what we should understand by the words "coordinates of the force system." We establish in this respect the following convention.

We introduce the values of $x', \ldots, r$ in terms of $\xi', \ldots, \vartheta'$ into the expression (2) for the work, and order the expression according to the latter quantities. *We then define the quantities that appear as the factors of ξ', η', ζ', φ', ψ', ϑ' to be the coordinates of the force system corresponding to the velocity coordinates ξ', η', ζ', φ', ψ', ϑ', respectively.* This definition of the force components stands in exact analogy to our definition of those words for a single mass particle. The force coordinate corresponding to ξ, for example, is the ratio of the work that our force system performs in the displacement $d\xi$ to this displacement.

The new force coordinates will obviously be linear functions of the old, and the transformation equations that lead from the latter to the

former appear entirely analogous to the transformation equations that express the old velocity coordinates in terms of the new; the coefficients of the horizontal and vertical ranks are merely interchanged with one another. We express this fact concisely when we say:

On the basis of our definition, the generalized force coordinates are always contragredient to the velocity coordinates.

We have already stated on page 79 the corresponding theorem, or, more correctly, the corresponding convention, for a single mass particle.

§3. The concept of the impulse for the generalized top. Relation between the impulse vector and the rotation vector. Connection to the expression for the *vis viva*

We now make the passage from statics to kinetics, and thus ask for the relation between the motion and the forces that cause the motion. The mass distribution of the body is now of decisive importance, so that we will from now on treat of the generalized and the symmetric top separately.

As in the case of the single mass particle, we place the *concept of the impulse* at the summit of kinetics. We illustrate this concept first for a freely moving rigid body, and proceed at once to a body fixed at one of its points.

The definition of the impulse is the following:

We consider the rigid body in an arbitrary state of motion, and ask for any system of impact forces that is capable of transforming the body instantaneously in its instantaneous position from rest to the considered state of motion. *This or any equivalent system of impact forces is called the impulse of the body.*

With consideration of the investigations of the preceding sections, we can immediately state the following theorems:

The impulse of a freely moving rigid body is a combination of a pushing-impact and a turning-impact; it can be conceived concisely as a screw.

and

The impulse of a rigid body with a fixed support point O is a single turning-impact; we can visualize it by the simple image of a vector emanating from O.

Remaining in the latter case, we place the kinematic vector of the rotational velocity into consideration alongside the static vector of the impulse.

A first exercise in the kinetics of the top will be to establish the mutual dependence of these two vectors.

To this end, we consider both the velocities and the impulses of all the individual mass elements of which the top is made.

We take the axis of the rotational velocity as the first coordinate axis of a rectangular coordinate frame XYZ, which has an unchanging position with respect to the body, and whose origin coincides with the support point. We denote the components of the impulse vector with respect to the coordinate axes by L, M, N, and those of the rotation vector, as earlier, by p, q, r. By assumption, the only nonzero value among the latter is p.

We now consider any element P of the body with mass dm. By virtue of the rotation along the X-axis, our element has a linear velocity

$$v = p\sqrt{Y^2 + Z^2}.$$

The impact force that is required to produce this velocity instantaneously has the magnitude $v\,dm$; its components with respect to the coordinate axes are, as one easily recognizes,

$$0, \quad -pZ\,dm, \quad pY\,dm.$$

An impact force of this form acts on our body for each distinguishable element P. The turning-force that corresponds to the system of these impact forces is then our impulse. Its components are calculated, according to the analytic rule of page 85, as

$$\text{(1)} \quad L = p\int(Y^2 + Z^2)\,dm, \quad M = -p\int YX\,dm, \quad N = -p\int ZX\,dm,$$

where the integral extends over the total mass of the body. The given expressions show immediately that the impulse vector generally differs in direction from the vector of the rotational velocity; while, according to assumption, the vector of the rotational velocity falls on the X-axis, the vector of the impulse also has components in the directions of the Y- and Z-axes.

If we assume that the instantaneous rotation occurs about the Y- or Z-axis, then we obviously have, by a cyclical interchange in a completely corresponding way, the following values for the components of the associated impulses:

$$(1') \quad L = -q\int XY\,dm, \quad M = q\int (Z^2+X^2)\,dm, \quad N = -q\int ZY\,dm,$$

and

$$(1'') \quad L = -r\int XZ\,dm, \quad M = -r\int YZ\,dm, \quad N = r\int (X^2+Y^2)\,dm.$$

But the impulse for a general position of the rotation vector now follows immediately from the preceding equations. As we know, rotational velocities, as well as turning-forces (impulses), compose as vectors; that is, their components simply add. To a rotation (p, q, r) about the axis $p:q:r$ there thus corresponds an impulse whose components are equal to the respective sums of the impulse components calculated in equations (1), (1′), and (1″). The associated impulse is thus

$$(2) \quad \begin{cases} L = p\int (Y^2+Z^2)\,dm - q\int XY\,dm - r\int XZ\,dm, \\ M = -p\int YX\,dm + q\int (Z^2+X^2)\,dm - r\int YZ\,dm, \\ N = -p\int ZX\,dm - q\int ZY\,dm + r\int (X^2+Y^2)\,dm. \end{cases}$$

The preceding equations immediately assume a very clear form if we introduce the quadratic form of the velocity components

$$(3) \quad T = \frac{1}{2}\left\{p^2\int (Y^2+Z^2)\,dm + q^2\int (Z^2+X^2)\,dm + r^2\int (X^2+Y^2)\,dm \right.$$
$$\left. - 2qr\int YZ\,dm - 2rp\int ZX\,dm - 2pq\int XY\,dm\right\}.$$

Namely, there then follow simply

$$(4) \qquad L = \frac{\partial T}{\partial p}, \quad M = \frac{\partial T}{\partial q}, \quad N = \frac{\partial T}{\partial r}.$$

We ask for the mechanical meaning of our quadratic form T. It appears that T *is the* vis viva *of the top; that is, the work which the impulse performs in the generation of the instantaneous state of motion.*

In fact, we calculate this work if we first compose the individual work that the individual applied impulses perform on the mass elements of the body. According to page 75, the work that is performed on the mass element dm in the generation of the velocity (x', y', z') is

$$dA = \frac{1}{2}(x'^2 + y'^2 + z'^2)\,dm.$$

The total work is calculated from this by integration over the total mass of the body; it becomes

$$\frac{1}{2}\int (x'^2 + y'^2 + z'^2)\, dm.$$

We reformulate this expression by introducing the rotational velocity p, q, r. We have produced the necessary expressions for x', y', z' on page 41. We thus obtain

$$\frac{1}{2}\int \{(-Zq + Yr)^2 + (-Xr + Zp)^2 + (-Yp + Xq)^2\}\, dm.$$

But the working out of this expression directly yields the right-hand side of equation (3). We will therefore say:

The vis viva *of the top is a homogeneous quadratic function of the components of the rotation vector, with constant coefficients that depend only on the mass distribution of the body.*

After we have recognized the meaning of T, we can declare equations (4) to be the exact analogue of equations (7) on page 76 for a single mass particle.

We will rewrite the expression for the *vis viva* in a series of other interesting forms. We first remark that according to a well-known theorem for homogeneous functions,

$$T = \frac{1}{2}\left(p\frac{\partial T}{\partial p} + q\frac{\partial T}{\partial q} + r\frac{\partial T}{\partial r}\right).$$

With consideration of equations (4), we can write instead

$$T = \frac{1}{2}(pL + qM + rN). \tag{5}$$

We express this formula in words in the following way:

The vis viva *is equal to half the product of the magnitude of the impulse vector and the projection of the rotation vector onto the impulse vector (or also equal to half the product of the magnitude of the rotation vector and the projection of the impulse vector onto the rotation vector).*

In the language of vector analysis (cf. page 62) we can also say concisely:

The vis viva *is equal to half the scalar product of the impulse vector and the rotation vector.*

We may also develop the last formula directly from the consideration of the total system, without entering into the synthesis of the rigid body from its individual mass elements, if we carry over a consideration given on page 75 for a single mass particle directly to our case.

We begin from the work that an arbitrary continuous turning-force (D^x, D^y, D^z) performs on our top in a displacement $p\,dt$, $q\,dt$, $r\,dt$. This work, according to equation (2′) of page 91, is

$$dA = (D^x p + D^y q + D^z r)\,dt. \tag{6}$$

From this we derive the expression for the finite work that our turning-impacts L, M, N perform in the *generation* of the rotation p, q, r (that is, exactly the expression for the *vis viva*) by integration with respect to time in the following manner.

We can conceive our turning-impact L, M, N as a continuous turning-force of very large constant magnitude and very small application duration Δt. We can therefore set

$$L = \int_0^{\Delta t} D^x\,dt = D^x \Delta t,\ M = \int_0^{\Delta t} D^y\,dt = D^y \Delta t,\ N = \int_0^{\Delta t} D^z\,dt = D^z \Delta t. \tag{7}$$

At the beginning of the interval Δt, the rotational velocity of the body equals zero, and at the end of the interval Δt it equals (p, q, r). We must now assume that in the intermediate time the velocity increases *uniformly*, so that

$$\int_0^{\Delta t} p\,dt = \frac{1}{2} p\Delta t,\quad \int_0^{\Delta t} q\,dt = \frac{1}{2} q\Delta t,\quad \int_0^{\Delta t} r\,dt = \frac{1}{2} r\Delta t. \tag{8}$$

If we integrate the expression (6) for the work between $t = 0$ and $t = \Delta t$, then we obtain, with consideration of (7) and (8),

$$\left\{\begin{aligned} T = \int_0^{\Delta t} dA &= D^x \int_0^{\Delta t} p\,dt + D^y \int_0^{\Delta t} q\,dt + D^z \int_0^{\Delta t} r\,dt \\ &= \frac{1}{2}(D^x p + D^y q + D^z r)\,\Delta t \\ &= \frac{1}{2}(Lp + Mq + Nr). \end{aligned}\right. \tag{9}$$

We thus return directly to equation (5).

In equations (4), we have assumed T to be a function of the velocity coordinates p, q, r. But we can also calculate T as a function of the impulse coordinates. It is enough, for this purpose, to solve equations (2) for p, q r and enter the resulting values of the latter quantities into (5). From (2) there first result

$$\left\{\begin{aligned} p &= A_{11}L + A_{21}M + A_{31}N, \\ q &= A_{12}L + A_{22}M + A_{32}N, \\ r &= A_{13}L + A_{23}M + A_{33}N, \end{aligned}\right. \tag{2′}$$

where the A_{ik} denote the values of the subdeterminants of the coefficient schema in (2) divided by the determinant, and where $A_{ik} = A_{ki}$. With consideration of (5), we now obtain for T the expression

$$(3') \qquad T = \frac{1}{2}(A_{11}L^2 + 2A_{12}LM + \cdots + A_{33}N^2).$$

Conceived as a function of the impulse coordinates, T is again a homogeneous quadratic form with constant coefficients.

We will, further, form the partial differential quotients of this function with respect to L, M, N. These will obviously be equal to the right-hand sides of equations (2′), so that we find the relations

$$(4') \qquad p = \frac{\partial T}{\partial L}, \quad q = \frac{\partial T}{\partial M}, \quad r = \frac{\partial T}{\partial N}.$$

The equations (4′) *represent the solution of equations* (4) *written in a characteristically symmetric form.* It is emphasized that T is expressed above as a function of p, q, r, and now as a function of L, M, N.

Equations (4), or the equivalent equations (4′), yield the desired relation between the impulse vector and the rotation vector in the most general form. They represent the first and most important equations in the kinetics of the top. They have, moreover, exactly the same form as the analogous equations for a single mass particle (cf. page 76). We can, collecting both triplets of equations, repeat the statement of page 77:

The impulse (velocity) components are the partial differential coefficients of the vis viva *taken with respect to the velocity (impulse) components, where we must think of the* vis viva *expressed as a function of the velocity (impulse) components.*

We next bring the expression for the *vis viva* into relation with the *concept of the moment of inertia.* The coefficients of $\frac{1}{2}p^2$, $\frac{1}{2}q^2$, $\frac{1}{2}r^2$ in expression (3) are designated, as is well known, as the moments of inertia of the body about the axis of X, Y, Z, respectively. On the other hand, the coefficients of $-pq$, $-qr$, $-rp$ in the same expression are occasionally called the "products of inertia" (or also the "centrifugal moments"). Further, the moment of inertia M of the body about an arbitrary axis is defined by the equation

$$M = \int R^2\, dm,$$

where R denotes the distance of the element dm from the considered axis, and where the integral extends over the entire mass of the body.

But we arrive at the same integrals in the expression for the *vis viva.* We note that the linear velocity of an element dm of our body is equal to the product of the angular velocity of the body about the instantaneous rotation axis and the distance of the element from this axis; if we denote the former by Ω and the latter by R, then

$$x'^2 + y'^2 + z'^2 = \Omega^2 R^2.$$

Thus

$$(10) \qquad T = \frac{1}{2}\int (x'^2 + y'^2 + z'^2)\, dm = \frac{\Omega^2}{2}\int R^2\, dm = \frac{M}{2}\,\Omega^2.$$

We compare the expression $T = \dfrac{M}{2}\Omega^2$ for the *vis viva* of a rigid body with the formula $T = \dfrac{m}{2}v^2$ for an individual mass particle. We can then say:

The vis viva *of the top is calculated from the angular velocity and the moment of inertia corresponding to the instantaneous axis of rotation in exactly the same manner as the* vis viva *of a single particle is calculated from the velocity and the mass.*

We may use equation (10), further, to establish a general expression for M. If we denote the direction cosines of the instantaneous rotation axis $p:q:r$ with respect to the coordinate frame XYZ by α, β, γ, so that

$$\alpha = \frac{p}{\Omega}, \quad \beta = \frac{q}{\Omega}, \quad \gamma = \frac{r}{\Omega},$$

then there results from (10) and (3)

$$(11) \quad \begin{cases} M = \alpha^2\displaystyle\int (Y^2+Z^2)\, dm + \beta^2\int (Z^2+X^2)\, dm + \gamma^2\int (X^2+Y^2)\, dm \\ \qquad\qquad -2\beta\gamma\displaystyle\int YZ\, dm - 2\gamma\alpha\int ZX\, dm - 2\alpha\beta\int XY\, dm. \end{cases}$$

The moment of inertia about an arbitrary axis (α,β,γ) is therefore a homogeneous quadratic function of the direction cosines α, β, γ, and depends on these quantities in completely the same manner as $2T$ depends on the velocity components p, q, r.

We introduce next the concept of the *ellipsoid of inertia*, customary since Poinsot, by first laying off on the axis (α,β,γ) the segment $\varrho = \sqrt{1/M}$ as a radius vector. The endpoint of this segment has the coordinates

$$\xi = \alpha\varrho, \quad \eta = \beta\varrho, \quad \zeta = \gamma\varrho.$$

If we make the same construction for all possible axes (α,β,γ), then

there results a surface of the second degree, and, in fact, an ellipsoid that has the equation

$$1 = \xi^2 \int (Y^2 + Z^2)\, dm + \eta^2 \int (Z^2 + X^2)\, dm + \zeta^2 \int (X^2 + Y^2)\, dm$$
$$-2\eta\zeta \int YZ\, dm - 2\zeta\xi \int ZX\, dm - 2\xi\eta \int XY\, dm.$$

The three principal axes of this ellipsoid are the so-called *principal inertial axes.* If we imagine the coordinates X, Y, Z shifted to the principal inertial axes, then the products $\eta\zeta$, $\zeta\xi$, $\xi\eta$ in the equation for the ellipsoid of inertia must vanish. *The principal axes are therefore distinguished in that the products of inertia are equal to zero if the principal axes are used as the coordinate axes.* The equation for the ellipsoid of inertia in these coordinates takes the form

$$1 = A\xi^2 + B\eta^2 + C\zeta^2, \tag{12}$$

where the quantities

$$A = \int (Y^2 + Z^2)\, dm, \quad B = \int (Z^2 + X^2)\, dm, \quad C = \int (X^2 + Y^2)\, dm$$

are called the *principal moments of inertia* with respect to the support point O.

Not every ellipsoid, moreover, can be an ellipsoid of inertia. Namely, one easily recognizes from the given expressions for A, B, C that these quantities satisfy the inequalities

$$A < B + C, \quad B < C + A, \quad C < A + B,$$

inequalities that we summarize in the simplest way by the statement that A, B, C are the sides of a possible straight-line triangle. *Thus the only ellipsoids that correspond to ellipsoids of inertia of actual bodies are those for which the squares of the reciprocals of the major axes can form a triangle.*

The expression (11) for the moment of inertia about an arbitrary axis goes over, in our present choice of coordinates, to

$$M = A\alpha^2 + B\beta^2 + C\gamma^2.$$

But the *vis viva* of the body transforms in just the same way as this expression. According to (10), we obtain for the *vis viva*

$$T = \frac{1}{2}(Ap^2 + Bq^2 + Cr^2). \tag{13}$$

Finally, equations (2) also simplify substantially if we let the coordinate frame coincide with the "principal coordinate frame." *Namely, the impulse vector (L, M, N) that corresponds to an arbitrary rotation*

vector (p, q, r) *is now determined through the fundamental equations*

$$L = Ap, \quad M = Bq, \quad N = Cr. \tag{14}$$

In consequence, there results from (13) the following expression for the *vis viva* in terms of the impulse coordinates:

$$T = \frac{1}{2}\left(\frac{L^2}{A} + \frac{M^2}{B} + \frac{N^2}{C}\right). \tag{13$'$}$$

We are able to conclude from equation (14), just as from the earlier equation (1), that the rotation vector and the impulse vector form, in general, a nonzero angle with each other. In fact, in so far as the principal moments of inertia are all different from one another, the proportion

$$L : M : N = p : q : r$$

is fulfilled only if two components of the rotation vector (impulse vector) vanish. *The rotation vector and the impulse vector thus coincide only if one of the two vectors (and therefore simultaneously the other) lies on one of the three principal axes.*

The relation between our two vectors may be described in geometric form, finally, by a simple construction that was already given, in essence, by Poinsot.

We begin from the ellipsoid of inertia

$$A\xi^2 + B\eta^2 + C\zeta^2 = 1$$

and place through the endpoint of the rotation vector p, q, r the ellipsoid similar and similarly situated with the ellipsoid of inertia. This ellipsoid will have the equation

$$A\xi^2 + B\eta^2 + C\zeta^2 = Ap^2 + Bq^2 + Cr^2 = 2T.$$

At the endpoint of the rotation vector we then place the tangent plane

$$Ap\xi + Bq\eta + Cr\zeta = 2T$$

to this ellipsoid. The perpendicular from O to this plane has the direction

$$Ap : Bq : Cr = L : M : N;$$

that is, the direction of the impulse vector. The length of the perpendicular is

$$\frac{2T}{G},$$

where

$$G^2 = L^2 + M^2 + N^2$$

denotes the square of the length of the impulse vector. If the rotation

vector, and therefore also the magnitude T of the *vis viva*, are given to us, then the direction and the magnitude of the impulse are determined from our construction. In particular, we can state the theorem:

The direction of the impulse vector lies perpendicular to the plane that is conjugate to the rotation vector with respect to the ellipsoid of inertia.[59]

A completely analogous construction leads to the magnitude and direction of the rotation vector if the impulse vector is given. We place the plane normal to the impulse vector through its endpoint. The equation of this plane is

$$Ap\xi + Bq\eta + Cr\zeta = G^2.$$

Among the surfaces similar and similarly situated with the ellipsoid of inertia, there is one that is tangent to our plane. It is the ellipsoid

$$A\xi^2 + B\eta^2 + C\zeta^2 = \frac{G^4}{2T}.$$

The binding line of the tangent point with O then yields the direction of the rotation vector. We obtain the magnitude of the rotation vector if we compare any one linear dimension of the latter ellipsoid with the corresponding linear dimension of the ellipsoid of inertia. Two such lengths stand in the proportion $G^2 : \sqrt{2T}$. Since G now is given, the magnitude of T and thus also the length of the rotation vector are known. —

It would not be difficult to give the corresponding development more generally for the case of the freely moving rigid body. We can limit ourselves to writing the results for this case directly, since their derivation differs only slightly from the development above.

We denote the coordinates of the impulse screw by X, Y, Z, L, M, N and the coordinates of the motion screw, as earlier, by x', y', z', p, q, r. The former quantities are determined from the latter in the simplest way by means of the equations

$$\text{(16)} \qquad \begin{cases} X = \dfrac{\partial T}{\partial x'}, \quad Y = \dfrac{\partial T}{\partial y'}, \quad Z = \dfrac{\partial T}{\partial z'}, \\[2ex] L = \dfrac{\partial T}{\partial p}, \quad M = \dfrac{\partial T}{\partial q}, \quad N = \dfrac{\partial T}{\partial r}. \end{cases}$$

Here T is the expression for the *vis viva* written as a function of the velocity coordinates. To arrange this expression most conveniently, one places the reference point at the center of gravity, and lets the coordinate system fixed in the body coincide with the principal axes passing

through the center of gravity. The *vis viva* T then becomes simply

$$T = \frac{m}{2}(x'^2 + y'^2 + z'^2) + \frac{1}{2}(Ap^2 + Bq^2 + Cr^2).$$

Written in terms of the impulse coordinates, T takes the form

$$T = \frac{1}{2m}(X^2 + Y^2 + Z^2) + \frac{1}{2}\Big(\frac{L^2}{A} + \frac{M^2}{B} + \frac{N^2}{C}\Big).$$

From this one recognizes that the inversion of equations (16) is

$$(16') \qquad \begin{cases} x' = \dfrac{\partial T}{\partial X}, & y' = \dfrac{\partial T}{\partial Y}, \quad z' = \dfrac{\partial T}{\partial Z}, \\ p = \dfrac{\partial T}{\partial L}, & q = \dfrac{\partial T}{\partial M}, \quad r = \dfrac{\partial T}{\partial N}, \end{cases}$$

where the expression for T in terms of the coordinates of the impulse screw is used for the differentiation. These equations are, to be sure, derived here for the sake of simplicity only for a special position of the reference point and for a special choice of the coordinate axes XYZ. They are, however, independent of this choice, and hold just as generally as equations (16).

The first parts of equations (16) and (16′) (referring to the motion of the center of gravity) are exactly identical to equations (7) and (7′) of the first section, and the second parts of equations (16) and (16′) (referring to the motion about the center of gravity) are exactly identical to equations (4) and (4′) of this section. They represent the first and most important determining equations in the kinetics of a free rigid body.

The expressions given above for the *vis viva* immediately result in the equation

$$T = \frac{1}{2}(x'X + y'Y + z'Z + pL + qM + rN),$$

which naturally can again be derived directly from the expression (2) on page 90 for the work performed in an infinitesimal displacement. The parenthesis on the right-hand of this equation naturally has a simple geometric meaning that depends only on the nature of the two screws and their relative positions, but not on their absolute position in space, and will be called the *moment of the two screws upon each other.* In terms of the pitches h and h' of the two screws, the shortest distance Δ, the angle of inclination φ of the two screw axes, the magnitude Ω of the rotational velocity, and the magnitude S of the pushing-impulse, the moment is expressed as*):

$$\Omega S\{2\pi\Delta\sin\varphi + (h + h')\cos\varphi\}.$$

*) Cf. F. Klein, Math. Ann. Bd. II, p. 368. Ball (l. c.) designates the expression in question as the "virtual coefficient" of the screws.

The concept of the impulse for the generalized top is completely developed by Poinsot in his repeatedly cited work. The designation of Poinsot is, somewhat ceremoniously, *couple d'impulsion* (translated in the German edition of Schellbach as "the motion-stimulating force-pair"(!)).

§4. Transference of the preceding results to the special case of the symmetric top

We now proceed specifically to the symmetric top, and therefore assume that our body has rotational symmetry about the figure axis. The question now is the simplification of the preceding kinetic consideration that results from this assumption. The ellipsoid of inertia naturally has rotational symmetry about the figure axis in the same way as the mass distribution of the body. The ellipsoid of inertia thus becomes a surface of revolution. In addition to the figure axis, all axes in the equatorial plane of the top become principal axes of the ellipsoid and principal inertial axes of the body. All the principal moments of inertia corresponding to these principal axes, moreover, are equal to each other.

If we wish to refer the top to a principal inertial frame as a coordinate frame, then we need only place the Z-axis on the figure axis; the axes of X and Y then fall in the equatorial plane, and become principal inertial axes with equal principal moments of inertia. If we denote, as earlier, the principal moments of inertia by A, B, and C, then we have *the characteristic relation for the symmetric top*

$$A = B.$$

For the present choice of the coordinate system, the equation of the ellipsoid of inertia is thus

$$A(\xi^2 + \eta^2) + C\zeta^2 = 1;$$

the expression for the *vis viva* becomes

$$T = \frac{1}{2}\Big(A(p^2 + q^2) + Cr^2\Big),$$

and the relations between the components of the impulse vector and the rotation vector are written in their simplest form as

$$L = Ap, \quad M = Aq, \quad N = Cr.$$

We wish to return first to the generalization of the definition of the symmetric top that was mentioned already in the Introduction. *We*

will always call a rigid body with a fixed support point O a symmetric top (or simply a top) if two of the three principal moments of inertia through O are equal to each other and, moreover, the center of gravity lies on the axis of the third principal moment of inertia. Such a body will behave, with respect to all questions concerning the rotation about the point O under the influence of gravity, exactly like a body that has the previously assumed *geometric* rotational symmetry about the figure axis. We will likewise carry over the designations "figure axis" and "equatorial plane of the top" to the figure axis and equatorial plane of the ellipsoid of inertia of our generalized body. The equatorial plane is thus distinguished by the fact that its collected axes are principal inertial axes with the same principal moment of inertia A. We can say of such a body that it has not a *geometric* rotational symmetry, but rather a *mechanical rotational symmetry about the figure axis.*

We will distinguish, moreover, three subclasses of symmetric tops, according to whether the ellipsoid of inertia is a prolate ellipsoid of revolution, an oblate ellipsoid of revolution, or, in particular, a sphere. We thus speak of a *prolate top*, an *oblate top*, or a *spherical top.* The spherical top is distinguished, in particular, by the fact that any axis passing through O represents a principal inertial axis of the body. Since the principal axes of the ellipsoid of inertia are the reciprocal values of $\sqrt{A}$ and $\sqrt{C}$, the ellipsoid of inertia will be prolate when $A > C$ and oblate when $A < C$. In the limiting case $A = C$, the ellipsoid of inertia goes over into a sphere. Thus the conditions for the three cases are

$$\begin{aligned} \textit{prolate top} &: A > C, \\ \textit{oblate}\ \ \text{''}\ \ &: A < C, \\ \textit{spherical}\ \ \text{''}\ \ &: A = C. \end{aligned}$$

As examples of the three types of tops with *geometric* rotational symmetry, we can always take an ellipsoid of revolution filled with homogeneous mass, which accordingly is prolate, oblate, or a sphere. It is also easy, however, to construct examples of tops with only *mechanical* rotational symmetry. In fact, four mass particles of equal mass that form the corners of a square, and are imagined to be bound to each other by rigid massless rods, represent a *symmetric top with an oblate ellipsoid of inertia* that has only *mechanical* rotational symmetry. If we fix a fifth mass particle on the figure axis of this top (that is, on the normal erected from the midpoint O of the square), then we

obtain, depending on the distance of this particle from O and its mass, an oblate top or a spherical top with again only *mechanical* rotational symmetry. In particular, we emphasize here for later use that we can construct in this manner a *spherical top that has an arbitrarily given (positive or negative) gravitational turning-moment P, and whose center of gravity does not, therefore, coincide with O.* For this purpose, we can choose the masses of four formerly cited particles equal, for example, to 1 gr, and the side length of the square in whose corners the particles are fixed equal to 1 cm. For the fifth particle, we must then arrange that its mass m and its distance E from O are

$$m = \left(\frac{P}{g}\right)^2, \qquad E = \frac{g}{P},$$

where g is the acceleration of gravity.[60]

We now carry out the *P o i n s o t* construction, through which we clarify the relation between the impulse axis and the axis of rotation for the symmetric top. The simplification here, compared with the earlier case of the generalized top, is that we can carry out the construction in a plane; namely, the meridian plane passing through the instantaneous axis of rotation. A characteristic distinction between our three types of tops is thus made noticeable.

We imagine the rotation vector as somehow given. Through the axis OR of this vector we place the meridian plane FOR, which we will use in the following as the plane of the drawing. We draw the figure axis vertically upward. The tangent plane at the point R', the intersection of the rotation axis with the ellipsoid of inertia or one of the similar and similarly situated ellipsoids used on page 101, is perpendicular to the plane of the drawing, and thus the perpendicular from O to this plane lies in the plane of the drawing. Instead of the tangent plane, it is therefore enough to consider the tangent to the ellipsoid of inertia lying in our meridian plane. The situation in detail is as follows.

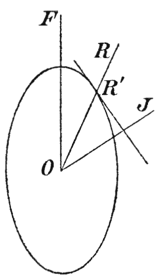


Fig. 13.

1. *The prolate top*, $A > C$. The perpendicular from O to the tangent at R' falls on the opposite side of the rotation axis as the figure axis (cf. Fig. 13). *For the prolate top, the rotation axis therefore lies between the impulse axis and the figure axis.*

2. *The oblate top*, $A < C$. The perpendicular from O to the tangent

at R' lies in the acute angle between the figure axis and the rotation axis (cf. Fig. 14). *For the oblate top, the impulse axis therefore lies between the rotation axis and the figure axis.*

3. *The spherical top*, $A = C$. Since the meridian intersection of the ellipsoid of inertia degenerates into a circle, the perpendicular from O

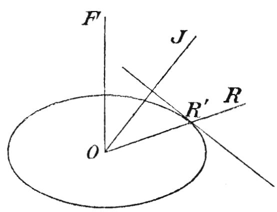


Fig. 14.

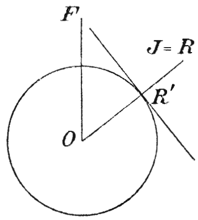


Fig. 15.

to the tangent through the contact point R' also passes through R' (cf. Fig. 15). *For the spherical top, the impulse axis and the rotation axis necessarily coincide.*

Only the spherical top behaves, so to speak, "*isotropically*"; that is, in the sense that the axis of the rotational motion coincides with the turning-force that generates the motion. The oblate and the prolate top exhibit the corresponding behavior only when the rotation is about the figure axis or about an axis of the equatorial plane, as is immediately evident from our construction. All these cases, moreover, are included in our general rule above, according to which the rotation vector and the impulse vector coincide in direction if and only if one of the two vectors lies along a principal inertial axis of the body. —

While we have thus far expressed the state of motion in terms of the components p, q, r, and have represented the impulse, correspondingly, in terms of the components L, M, N, we now wish to think of the instantaneous rotation of the symmetric top as given in terms of the changes in the Euler angles φ, ψ, ϑ, and ask for the "corresponding components of the impulse." The following considerations are also valid, moreover, for the generalized top. We give the development here first for the symmetric top only because the formulas for the generalized top become somewhat long.

We have already established on page 92 how we define, in general, the force coordinate corresponding to a velocity coordinate for a rigid body. (We refer also to the completely analogous consideration on page 78 for a single mass particle.) What was said for the force coordinates naturally holds equally well for the impulse coordinates; what was

developed for the freely moving rigid body carries over in an immediately more understandable manner to the top.

According to the noted rule, we proceed from the equations that express the old velocity coordinates p, q, r in terms of the new coordinates φ', ψ', ϑ'. These are the "kinematic" equations

$$(1)\qquad \begin{cases} p = \qquad\quad \psi' \sin\vartheta \sin\varphi + \vartheta' \cos\varphi, \\ q = \qquad\quad \psi' \sin\vartheta \cos\varphi - \vartheta' \sin\varphi, \\ r = \varphi' + \psi' \cos\vartheta \end{cases}$$

of page 45. As a result, the new impulse components, which we denote by $[\Phi]$, $[\Psi]$, $[\Theta]$, depend on the old components in the following way:

$$\begin{aligned} [\Phi] &= N, \\ [\Psi] &= L \sin\vartheta \sin\varphi + M \sin\vartheta \cos\varphi + N \cos\vartheta, \\ [\Theta] &= L \cos\varphi - M \sin\varphi. \end{aligned}$$

Here we replace L, M, N by their expressions in terms of φ', ψ', ϑ', which are given by equations (1), if we multiply those equations by A, A, and C, respectively. Thus we obtain

$$(2)\qquad \begin{cases} [\Phi] = C(\varphi' + \psi' \cos\vartheta), \\ [\Psi] = C \cos\vartheta\, \varphi' + (C \cos^2\vartheta + A \sin^2\vartheta)\psi', \\ [\Theta] = A\vartheta'. \end{cases}$$

We refer once more to the relation that holds between the impulse and velocity coordinates and the partial differential quotients of the *vis viva*. It is self-evident from our definition of the impulse coordinates that this relationship, which we know to be valid for the velocity coordinates p, q, r, remains valid if we introduce new coordinates φ', ψ', ϑ' that depend linearly on the old coordinates. But we may nevertheless convince ourselves of this as follows. According to equations (1), the expression for the *vis viva*

$$T = \frac{1}{2}\left(A(p^2 + q^2) + Cr^2\right)$$

is written in terms of the coordinates φ', ψ', ϑ' as

$$(3)\qquad T = \frac{1}{2}\left(A(\vartheta'^2 + \sin^2\vartheta \,.\, \psi'^2) + C(\varphi' + \cos\vartheta \,.\, \psi')^2\right).$$

But from this there result immediately, in analogy to equations (4) of the previous section,

$$(4)\qquad [\Phi] = \frac{\partial T}{\partial \varphi'}, \quad [\Psi] = \frac{\partial T}{\partial \psi'}, \quad [\Theta] = \frac{\partial T}{\partial \vartheta'}.$$

The relations analogous to equations $(4')$ are easily verified in the same way.

Finally, we ask for the *geometric meaning of our impulse coordinates* $[\Phi]$, $[\Psi]$, $[\Theta]$; we infer this from the geometric meaning of the expression $(2')$ on page 91 for the work of an infinitesimal displacement of our top.

Just as we could conceive the expression for the *vis viva* on page 96 as half the product of the length of the rotation vector and the projection of the impulse vector onto the rotation vector, we will now say that *the expression*

$$dA = (D^x p + D^y q + D^z r)\, dt$$

for the work that would be performed by an arbitrary turning-force D applied to our top in an infinitesimal displacement $(p, q, r)\, dt$ is equal, up to the factor dt, to the product of the magnitude of the turning velocity and the projection of the turning-force onto the axis of the turning velocity. If we now introduce our velocity coordinates φ', ψ', ϑ' instead of p, q, r, then the expression for the work retains, as we know, its earlier form. The preceding equation is thus transformed, if we denote the corresponding components of D by Φ, Ψ, Θ, into

$$dA = (\Phi\varphi' + \Psi\psi' + \Theta\vartheta')\, dt.$$

We now consider, in particular, an infinitesimal rotation for which $\psi' = \vartheta' = 0$, so that the corresponding work becomes equal to $\Phi\varphi'\, dt$. In this case, the rotation vector lies along the figure axis, since φ' denotes a rotation about the figure axis. From the geometric meaning of dA, it now follows immediately *that Φ denotes the perpendicular projection of the vector D onto the figure axis.* Further, the rotation vectors ψ' and ϑ' fall, respectively, in the directions of the vertical and the line of nodes. Thus it follows in the same way *that Ψ and Θ represent the perpendicular projections of the vector D onto the vertical and the line of nodes.* Exactly the same geometric meaning naturally belongs to our impulse coordinates $[\Phi]$, $[\Psi]$, $[\Theta]$. *These quantities are equal, respectively, to the perpendicular projections of the impulse vector onto the figure axis, the vertical, and the line of nodes.*

We can, without further elaboration, generalize the result of the last consideration by saying: *if, on the basis of any three skew-angular axes, we decompose the rotation vector into components parallel to these axes, then we obtain the corresponding decomposition of the force or impulse vector if we project perpendicularly to those axes.*

§5. The two fundamental theorems on the behavior of the impulse vector in the course of the motion

While we have oriented ourselves thus far with respect to the impulse that corresponds to an instantaneous state of motion of a body, or, equivalently, to the instantaneous state of motion resulting from a given impulse, it will be our next exercise to investigate the course of the motion in time. The previously discussed relation between the impulse vector and the instantaneous rotation vector is completely independent of the external circumstances under which the motion proceeds; that is, of the continuous forces that act on the rigid body. Our further considerations, however, will be determined in an essential way by these external forces. In this respect, we assume, on the one hand, that our body is entirely free of external forces, including, in particular, the force of gravity. On the other hand, we will admit of arbitrary continuous forces. We argue first for the freely moving rigid body. In doing so, we place the *investigation of the impulse* before the *consideration of the state of motion*. We therefore ask, in the first place: *how is the impulse of our body changed for force-free motion?* We derive the change of the state of motion from the behavior of the impulse only later. The answer to our question is simply the following:

The impulse is completely unchanged; it remains constant in space during the motion.

We establish this fundamental theorem in the most elementary manner from the kinetics of individual mass particles as follows.

We begin with a single mass element P, which moves freely and is subjected to no external forces. As we know, the direction and the magnitude of the impulse of such an element remain constant in space. (Galileo's law of inertia.)

We now consider two mass elements P and P' that are rigidly bound and are free, moreover, of external forces. We replace the action of the rigid binding dynamically by forces; in particular, we have a force directed toward P' at the point P and a force of the same magnitude directed toward P at the point P'. (Newton's lex tertia.) The common magnitude of the two forces depends on the requirements of the rigid binding, and modifies, in magnitude and direction, the directly applicable individual impulses of P and P'. If we include these forces — we can call them reaction forces — we may then treat of our two mass particles as freely moving particles. Now the individual impulses of P and P'

change continuously as a result of the presence of the reaction forces; the impulses add geometrically with the infinitesimal impacts corresponding to the reaction forces (Newton's lex secunda). It is different with the total impulse of the system formed by the two mass particles. We construct this total impulse by composing the individual impulses of the two particles according to the rules of elementary statics. In this construction, however, the two oppositely directed equal reaction forces obviously cancel at each moment. The resulting impact-screw therefore behaves exactly as if our reaction forces did not exist and our mass particles were free. *The impulse thus remains constant for the entire duration of the motion.*

A system of three rigidly bound mass particles that are free of external forces behaves no differently. Here we must consider not one but rather three pairs of equal and opposite reaction forces that act on the sides of the triangle formed by our particles. The individual impulses of the system particles are again changed by these reaction forces. In the construction of the impact-screw corresponding to the entire system, however, the reaction forces do not come into consideration; this screw behaves exactly as if our three mass particles moved freely according to the Galilean law of inertia.

The same deliberation carries over immediately to the case of arbitrarily many bound mass particles, and, further, to a spatially continuous mass system that is free of external forces. It holds even for the more general case of a nonrigid system whose particles are subjected only to internal forces that satisfy the principle of action and reaction; the result is valid, for example, for an elastic body, for the planetary system, or for a fluid quantum.

For the case of a freely moving rigid body, we state the result here explicitly as the first of the theorems governing the course of the motion:

T h e o r e m I: *The impulse-screw of a rigid body remains constant in space for force-free motion.*

It will be useful to carry over this theorem into the language of ordinary analytic mechanics. To this end, we calculate the components of the impulse-screw according to the rule on page 85; for P_i^x, P_i^y, P_i^z, we must take the components of the individual impulses of all the mass particles that constitute the body; we therefore set

$$P_i^x = x_i' \Delta m_i, \quad P_i^y = y_i' \Delta m_i, \quad P_i^z = z_i' \Delta m_i$$

and must finally go over from summation to integration. We thus obtain

$$S^x = \int x' \, dm = c',$$

$$S^y = \int y' \, dm = c'',$$

$$S^z = \int z' \, dm = c'''.$$

The integrals appearing here are nothing other than the velocities of the center of gravity (multiplied by m). *Our theorem thus states that the velocity of the center of gravity is constant; the theorem is identical with the simplest case of the so-called center of gravity theorem.*

In the same manner, we calculate the turning-moments D^x, D^y, D^z of our screw about O. According to the same rule as above, there follow

$$\begin{aligned} D^x &= \int (z'y - y'z) \, dm = c^{IV}, \\ D^y &= \int (x'z - z'x) \, dm = c^{V}, \\ D^z &= \int (y'x - x'y) \, dm = c^{VI}. \end{aligned}$$

These equations are also well known to us from ordinary mechanics; they are simply the so-called area theorems.[61] *The second part of our theorem regarding the turning components is therefore identical with the area theorems,* which, as is well known, apply to the free motion of a rigid body.

One notes, in particular, that our geometric consideration implies certain simple integrations that one is forced to carry out in the analytic derivation. The equivalent of the integration process is the recognition that the infinitesimal additional impulses which originate from the reaction forces are mutually annihilated in the formation of the total impulse.

We could have defined the impulse of the body directly in terms of the constants of the center of gravity and the area theorems; the opposite path followed here, however, appears to us more instructive. Our derivation forced us to return to the true root of the theorems, the

mechanical principles that are otherwise easily hidden behind formulas; it leaves, we believe, nothing more to be desired in transparency. It corresponds completely, moreover, to the tendencies of P o i n s o t, who gives, for his part, a much less simple proof*). In contrast, a consideration entirely analogous to the above is found in a beautiful work of R. B. H a y w a r d**), to which we will make further reference.[62]

The transference of our theorem to a body with a fixed support is now given immediately. We must now decompose our impulse screw at the fixed point O into a pushing-impact S and a turning-impact D. The former will be canceled by the fastening of the body; that is, by the reaction force at the support point. There remains only the turning-impact, and this, for the top, is what we called the impulse. There follows again, in analogy to the above,

T h e o r e m I a: *The impulse vector of a body supported at one point remains constant in magnitude and direction in space for force-free motion.*

This theorem is identical with the statement that the *area theorems* of the previous case remain valid, while the *center of gravity theorems* obviously become invalid. At the same time, the theorem appears as the exact analogue of the Galilean law of inertia, if we give the latter the form of Theorem I of page 74. —

It is now assumed that *arbitrary continuously applied external forces* act at the points of our body; the body will again be taken first as freely moving. Its impulse will no longer remain constant, and the question will be to determine the manner in which it is changed.

We arrange this consideration as above. If the body consists of a single particle, then its impulse and the impulses of the infinitesimal impacts of the external forces compose successively at each moment according to the parallelogram rule (Newton's lex secunda). If the body consists of two particle at an invariable distance of separation, then the individual impulses of these two particles will be changed by the external forces as well as by the reaction forces acting between them, by which we replace the rigid binding. If we consider, however, the total impulse of the system formed by the two particles, then the reaction forces cancel. The impact-screw in question therefore consists of a constant part that corresponds to the original impulse of our two

*) Théorie nouvelle ..., Kap. II, §5.

**) On a direct method of estimating velocities with respect to axes moveable in space. Cambridge Phil. Transact. Vol. X, 1856.

particles, and a variable part, which is caused only by the additional impacts corresponding to the external forces in the considered time period. We can first compose these latter impacts at each moment into an infinitesimal pushing-impact dS and a turning-impact dD with respect to a reference point O, and then combine dS and dD with the corresponding components S and D of the current impulse-screw according to the parallelogram construction. This procedure justifies the following theorem, which we immediately generalize to a rigid system of arbitrarily many particles and, further, to a rigid body of continuous mass distribution (not to mention generalized systems with only internal forces):

Theorem II: *The impulse screw of a freely moving rigid body on which arbitrary external forces act is changed during the motion so that it is composed at each moment according to the rules of statics with the infinitesimal impact screw caused by the external forces.*

This theorem corresponds to the cessation of validity of the simple center of gravity and area theorems for motion influenced by external forces. For one of the simple area or center of gravity theorems to remain valid, the system of external forces must fulfill a special condition; it must, as one says according to the precedent of Lie, admit of a certain infinitesimal transformation, namely an infinitesimal rotation or translation. In this case, a turning or pushing component of the force-screw corresponding to the external forces would vanish; at the same time, our geometric construction would immediately imply that a component of the impulse-screw would remain constant during the motion.

One notices the characteristic inappropriateness and asymmetric manner of notation that is used in analytic mechanics with respect to the center of mass and area theorems. One speaks of the existence of an area theorem only when the turning-moment of the external forces about an axis is equal to zero; that is, only when the turning part of the impulse has an invariable component in time. In contrast, one speaks of the center of gravity theorem also when external forces are applied, and therefore if the pushing part of the impulse changes arbitrarily during the course of the motion. This discrepancy arises essentially because the concept of the impulse, which is bound organically to the area and center of mass theorems, is not considered in the usual

presentations. The natural use of language would obviously be to conceive the words area theorem just as generally as the words center of gravity theorem, and therefore to understand the area theorem to mean that the turning-impact of the impulse adds geometrically with the successive turning-impacts of the exterior forces. In the case of constant turning- and pushing-impacts, one would then speak, as above, of the "simple" area and center of gravity theorems.

We make once more the passage to a body with a fixed support point, for which the constituent part S of the impulse will be canceled by the reaction force at O. Our deliberation above then leads to

Theorem IIa: *The impulse vector of a top on which arbitrary continuous external forces act is altered at each moment so that its change in direction and magnitude is equal to the infinitesimal turning-impact caused by the external forces.*

This theorem*) coincides in form precisely with the second Newtonian axiom, in so far as we state the latter as in Theorem II of page 74.

§6. The theorem of the *vis viva*

The preceding impulse theorems, together with our earlier relation between the impulse vector and the rotation vector, determine the motion of the top just as completely as the Newtonian axioms, to which they correspond exactly in form and content, govern the mechanics of a single mass particle. In fact, the successive changes of the impulse in space are established by our previous considerations. From these changes in impulse, however, follow the position of the rotation vector in the body, and therefore also the instantaneous motion, by virtue of the results of §3.

It will, therefore, no longer be necessary to return to the synthesis of the body from its individual mass elements and to follow (by means of the Newtonian axioms) the motion of the latter from the standpoint of particle mechanics. If we later do this on occasion, as, for example, at the end of this section, it is only on secondary and didactic grounds, since particle mechanics is especially familiar through general use.

*) In a monograph on the top: *A. de Saint-Germain, Résumé de la théorie du movement d'un solide autor d'un point fixe*, Paris 1887, this theorem is attributed in error to Résal. The first volume of the *Traité de Mécanique général* by Résal, in which the theorem in question appears on page 247, first appeared in 1873, while our theorem (compare, for example, the year of the Hayward treatise) is certainly much older.

The preceding impulse theorems must also include, in so far as the top is concerned, the theorem of the *vis viva*. In fact, this theorem, as we will immediately show, is only a corollary of our impulse theorems.

We first assume that no external forces act on our body, naturally disregarding the reaction force at the support point and such forces that are canceled by this reaction force.

The expression for twice the *vis viva*

$$Lp + Mq + Nr \tag{1}$$

signifies geometrically, as mentioned on page 96, the scalar product of the impulse and rotation vectors, and has, as such, a value that depends only on the magnitudes and the relative positions of the two vectors, and not on their positions in space.

We temporarily consider a uniform rotation of the body about the axis $p:q:r$, which we imagine fixed in the body and therefore also in space, while at the same time the vector (L, M, N) will be imagined as fixed in space. The change in the above scalar product due to this motion, that is, the quantity

$$p\,dL + q\,dM + r\,dN,$$

is equal to zero, since the magnitudes and the relative positions of our two vectors are not changed.[63]

The actually occurring force-free motion can, however, in so far as the behavior of the impulse vector and the motion of the body are concerned, be identified at each moment, in the first approximation, with a motion of the assumed nature. Thus it is also valid for the actual motion that

$$p\,dL + q\,dM + r\,dN = 0. \tag{2}$$

We next remark that for the actual motion, according to equation (4′) of §3,

$$p = \frac{\partial T}{\partial L}, \quad q = \frac{\partial T}{\partial M}, \quad r = \frac{\partial T}{\partial N},$$

where T is understood as the expression for the *vis viva* written in terms of the impulse coordinates. Thus the left-hand side of equation (2) is transformed into the perfect differential of the *vis viva*. We therefore obtain the equation $dT = 0$, or integrated $T = h$, which yields the theorem:

For force-free motion of the top, the vis viva *of the body is not changed.*

The *conservation of the* vis viva thus follows immediately, as we see, from the *conservation of the impulse.*

Now let, on the other hand, arbitrary external forces act on our top. We compose these forces, with respect to the support point O, into a pushing-force and a turning-force. The former we can neglect, and we denote the components of the latter, in the coordinate system to which L, M, N are also referred, by $\mathsf{\Lambda}$, M, N. In the time element dt, the impulse vector in space now undergoes the displacement $\mathsf{\Lambda}\,dt$, $\mathsf{M}\,dt$, $\mathsf{N}\,dt$ (Theorem IIa of the preceding §). The impulse vector, therefore, does not retain its magnitude and position in space. We must, at each moment, make in reverse the displacement of the endpoint of the impulse vector caused by the external forces in order to obtain a point fixed in space. The displacement of this point relative to the body is, resolved into components,

$$dL - \mathsf{\Lambda}\,dt, \quad dM - \mathsf{M}\,dt, \quad dN - \mathsf{N}\,dt.$$

Its binding segment with O yields a vector that has the same magnitude and relative position with respect to the rotation vector at the end of the time interval dt as the impulse vector had at the beginning of the time interval.

The same thus holds for this vector as for the impulse vector itself in force-free motion. Equation (2) is, therefore, now to be replaced by the equation

$$(3) \qquad p\,dL + q\,dM + r\,dN = (\mathsf{\Lambda}p + \mathsf{M}q + \mathsf{N}r)\,dt.$$

The left side of this equation is again the perfect differential of the *vis viva* (dT); the right side represents (according to equation $(2')$ of page 91) the work performed during the time interval dt by the external forces. We thus have the theorem:

For motion of the top influenced by external forces, the vis viva *is altered at each moment so that its change is equal to the infinitesimal work performed by the external forces* $(dT = dA)$.

It can occur, in particular, that the finite work which the external forces perform on our body, when we bring the body from a fixed initial position into any new position, depends only on this end position, and not on the intermediate positions passed through in the motion. As is well known, one calls the negative value of this work the *potential energy* U, and correspondingly designates T as the kinetic energy and $T + U$ as the total energy of the body. Then dA is the perfect differential of the

function $-U$. As a result, the preceding theorem takes the simpler form $dT = -dU$, or $T + U = h$, and can be stated in the following way:

If the external forces that influence the motion have a "potential," the total energy of the body does not change during the motion.

This *theorem of the change of the kinetic energy or the conservation of the total energy* is therefore also, as we see, a simple consequence of our *theorem of the change of the impulse.*

We prove in a corresponding manner the theorem of the *vis viva* for the motion of a free rigid body.

Here we must replace the expression (1) by the expression

$$Xx' + Yy' + Zz' + Lp + Mq + Nr, \tag{4}$$

in which x', y', z', p, q, r denote the coordinates of the instantaneous motion-screw and X, Y, Z, L, M, N denote the coordinates of the impulse-screw. This expression, as emphasized on page 103, has a geometric meaning that is independent of the position of the two screws in space, and depends only on their pitches and their relative positions.

Force-free motion is treated first. We carry out the infinitesimal screw $(x', y', z', p, q, r)\,dt$ and consider the relative motion of the impulse screw with respect to the body. We imagine the motion-screw fixed in the body and therefore also in space; the impulse-screw, which according to the previous section is fixed in space, will thus perform a screw motion about the motion-screw, so that its relative position with respect to the motion-screw and its pitch are not changed. If dX, dY, dZ, dL, dM, dN denote the relative coordinate changes of the impulse, then there holds for the motion considered here, and also for the actual motion,

$$x'\,dX + y'\,dY + z'\,dZ + p\,dL + q\,dM + r\,dN = 0. \tag{5}$$

But the left side is, in consequence of equation $(16')$ of page 103, the perfect differential dT of the *vis viva*; we thus have $dT = 0$, or $T = h$.

The vis viva *again remains unchanged for force-free motion of a rigid body.*

For the case in which arbitrary external forces influence the motion of the rigid body, the consideration is generalized immediately in the following evident manner.

We first compose the exterior forces, with respect to the reference point, into a pushing-force $(\Xi, \mathsf{H}, \mathsf{Z})$ and a turning-force $(\Lambda, \mathsf{M}, \mathsf{N})$. The changes of the impulse coordinates relative to space during the time

element dt are equal, according to Theorem II of the previous section, to $\Xi\,dt$, $\mathsf{H}\,dt$, $\mathsf{Z}\,dt$, $\Lambda\,dt$, $\mathsf{M}\,dt$, $\mathsf{N}\,dt$. The changes dX, dY, dZ, dL, dM, dN of the impulse coordinates relative to the body thus come only in part from the infinitesimal screw $(x', y', z', p, q, r)\,dt$; another part will be caused by the external forces. We must make the changes corresponding to the latter in reverse to obtain a screw that lies relative to the motion-screw at the end of the time interval dt just as the impulse-screw lay at the beginning of the infinitesimal motion. In other words, we must replace the quantities dX, …, dN in equation (5) by

$$dX - \Xi\,dt,\; dY - \mathsf{H}\,dt,\; dZ - \mathsf{Z}\,dt,\; dL - \Lambda\,dt,\; dM - \mathsf{M}\,dt,\; dN - \mathsf{N}\,dt.$$

Thus there follows

$$\begin{aligned} &x'\,dX + y'\,dY + z'\,dZ + p\,dL + q\,dM + r\,dN \\ &= (\Xi\,x' + \mathsf{H}\,y' + \mathsf{Z}\,z' + \Lambda\,p + \mathsf{M}\,q + \mathsf{N}\,r)\,dt. \end{aligned} \tag{6}$$

The left side, however, is the perfect differential of the *vis viva*, and the right side denotes, according to equation (2) of page 90, the work performed by the external forces. We thus have $dT = dA$:

The change of the vis viva *is equal at each moment to the work performed by the external forces.*

It is perhaps useful to carry out the proof of this theorem yet again according to the method of the previous section, in which we imagine our body resolved into its individual mass elements. It is enough, for this purpose, to consider a system of two rigidly bound mass elements.

We remark in advance that the change of the *vis viva* of a single mass particle is equal, on the basis of the formula

$$dT = x'd[X] + y'd[Y] + z'd[Z],$$

to the scalar product of the velocity vector (x', y', z') with the change of the impulse vector $([X], [Y], [Z])$.

At each of our two rigidly bound particles 1 and 2, we imagine the vectors of the individual impulses 1 and 2, which coincide in direction with the velocity vectors 1 and 2; we imagine as well the reaction forces 1 and 2, which replace the rigid binding of the points and act along the binding line of the points. External forces may not be present.

An evident consequence of the rigid binding is that the projection of the velocity vector 1 on the binding line is equal to that of the velocity

vector 2. In place of this, we can say, on the basis of the Newtonian lex tertia:

The sum of the scalar products of the velocity vectors with the corresponding reaction forces is equal to zero.

In fact, the equality of the considered projections is transformed into the negative equality of the considered scalar products because of the opposing sense of the two reaction forces. Now the reaction forces, however, are proportional in magnitude and direction to the changes of the individual impulses (Newton's lex secunda). *Therefore the sum of the scalar products of the changes of the individual impulses with the individual velocity vectors will also equal zero.*

According to the prefatory remark, however, the two terms of this sum equal the respective changes of the *vis viva* of our two mass particles. The sum itself is therefore equal to the change of the *vis viva* of the system.

Thus the vis viva *of our system remains constant, just as for a single mass particle that moves according to the Galilean law of inertia.*

The generalization of our deliberation to the case in which external forces act, or to that of arbitrarily many particles bound to form a rigid system, as well as the specialization to the case of the top, is so simple that we can pass over it.

We remark, further, that the analytic proof which is usually given for the theorem of the *vis viva* runs precisely parallel to the preceding geometric proof. Indeed, the last consideration, in which we returned to the individual mass elements of the rigid body, corresponds in analytic mechanics to taking as a basis the differential equations in the form of the so-called *Lagrange equations of the first kind*, while the earlier consideration of the total system corresponds to the standpoint of the so-called *generalized Lagrange equations (the equations of the second kind).* We will enter more deeply into both systems of equations in the following chapter (cf. §3).

§7. Geometric treatment of force-free motion of the top

We will now use the preceding general theorems to provide a clear geometric image of the motion of the top in the simplest conceivable case. We assume that no external forces act on the top. To specifically eliminate the effect of gravity, we imagine the top supported at its center of gravity.

The geometric theory of this motion, which was first given by P o i n s o t, can now be written down immediately.

We consider, in the first place, that the impulse vector of the top remains constant in space for force-free motion. We imagine, once and for all, that this vector is erected vertically upward from O. The magnitude and direction of this vector are naturally given by the initial impact through which our body has been set in motion.

In the second place, we consider that the *vis viva* of the body also remains constant for force-free motion. We give this fact a twofold geometric expression.

On the one hand, the *vis viva* denotes one-half the product of the magnitude of the impulse vector with the projection of the rotation vector onto the impulse vector. From the constancy of the impulse and the constancy of the *vis viva* taken together, it follows that the projection of the rotation vector onto the impulse vector has an invariable length. The magnitude of this projection depends again on the nature of the original impulse. *We therefore have a plane e, fixed in space and perpendicular to the impulse axis, that yields a locus for the endpoint of the rotation vector with respect to its position in space.*

A further geometric meaning of the theorem of the *vis viva* results from the expression

$$T = \frac{1}{2}(Ap^2 + Bq^2 + Cr^2).$$

The endpoint (p, q, r) of the rotation vector thus lies on an ellipsoid rigidly bound to the top, which is similar and similarly situated with the ellipsoid of inertia. The constancy of the *vis viva* now means that this ellipsoid continuously retains it dimensions during the motion. *We therefore have, in addition, an ellipsoid E, fixed in the body, that yields a locus for the endpoint of the rotation vector with respect to its position in the body.*

We consider, finally, the relation between the impulse vector and the rotation vector. This relation can be expressed by means of the Poinsot construction of page 101, so that the tangent plane of the ellipsoid E at the endpoint of the rotation vector stands perpendicular to the axis of the impulse. This tangent plane is thus, first of all, permanently parallel to our plane e; since the plane e, moreover, constantly passes through the endpoint of the rotation vector, the tangent plane coincides directly with the plane e. In other words:

Our ellipsoid E is constantly tangent to our plane e during the motion.

Now the polhode curve runs somehow in the ellipsoid E, and the herpolhode curve runs somehow in the plane e. Since the two curves roll on one another in the motion, our ellipsoid E also rolls without sliding on our plane e during the motion. *Thus we can emulate the completely kinetically determined motion in a purely kinematic way if we let an ellipsoid, described with O as its midpoint and fixed in the body, roll without sliding on a plane fixed in space.*[64]

We owe this beautiful and transparent image for the motion of the force-free top, as is well known, to the investigations of Poinsot. The motion in question will therefore be designated concisely as *Poinsot motion.* If, following the precedent of Poinsot, we thus visualize the motion by the rolling of an ellipsoid on a plane, we place ourselves, curiously, in a certain opposition to the general Poinsot theory of rotation. According to this general theory, we should, in the first place, make clear to ourselves the forms of the rolling cones and seek from these to acquire a representation of the motion. But the forms of these cones are rather complicated even in the preceding simple example, as we will show below for the herpolhode cone, and the differing construction above is much more transparent. In more difficult cases (with the addition of gravity), we can expect even less to manage with the discussion of the rolling cones alone.

To treat only of the successive positions of the body in space, our picture of the motion above is fully sufficient. If we wish, however, to give expression also to the velocity of the actual motion in our kinematic picture, then we must add a further specification regarding the manner of the rolling:

The velocity of rolling should be measured so that the rotation of the ellipsoid, which naturally takes place about the radius erected from O to the momentary tangent point of E and e, is equal to this radius.

The way in which this condition is also to be realized purely kinematically was first shown by Sylvester*). We cannot enter into this here.[65]

We now complete our representation of the Poinsot motion by studying, in detail, the behavior of the different geometric elements of the motion.

*) Cf. Sylvester: On the motion of a rigid body etc., London R. S. Phil. Transactions 1866.

The course of the motion with respect to the body is treated first. In this respect we wish to know, above all, *the curve that the endpoint of the impulse vector describes relative to the body.* We will occasionally call this curve, for want of a better expression, the "impulse curve."

In a single time element of the relative motion under discussion, the endpoint of the impulse vector always moves in a circular arc around the instantaneous rotation axis. Its relative displacement with respect to the body is therefore perpendicular to the instantaneous rotation axis. The direction of this displacement with respect to the XYZ frame is determined by the ratios of the coordinate changes $dL : dM : dN$, and the direction of the rotation vector is determined through the ratios $p:q:r$. Thus the equation

$$p\,dL + q\,dM + r\,dN = 0$$

is satisfied.

On the basis of the general relation between the impulse vector and the rotation vector, we can also write our equation as

$$\frac{L\,dL}{A} + \frac{M\,dM}{B} + \frac{N\,dN}{C} = 0.$$

Through integration there follows

$$\frac{1}{2}\left(\frac{L^2}{A} + \frac{M^2}{B} + \frac{N^2}{C}\right) = h. \tag{1}$$

Equation (1) is naturally equivalent to the equation of the *vis viva*, since our deliberation given in the previous section for the proof of the theorem of the *vis viva* is only repeated in a somewhat more special setting.

It is clear, further, that the endpoint of the impulse vector, because of its constant length G, must always lie on a sphere of radius G. We thus have

$$L^2 + M^2 + N^2 = G^2. \tag{2}$$

The desired curve is determined by equations (1) and (2). We can say:

The path that the endpoint of the impulse vector describes with respect to the body is a spherical curve; it is given by the intersection of the ellipsoid (1) *with the sphere* (2).

The form of the curve is sketched in Figures 18, 19, and 20 (see the following section) for a few characteristic cases.

At the same time, the curve that the endpoint of the rotation vector describes with respect to the body (that is, the polhode curve)

is also determined. We obtain this curve from the just derived curve of the impulse through a simple deformation with respect to the principal axes of the body. It also lies simultaneously on two surfaces of the second degree, the first of which is already known to us as our ellipsoid E. Namely, it follows from (1) and (2), with consideration of the relation between the impulse vector and the rotation vector, that

$$\frac{1}{2}(Ap^2 + Bq^2 + Cr^2) = h \tag{3}$$

and

$$A^2p^2 + B^2q^2 + C^2r^2 = G^2. \tag{4}$$

The polhode curve is therefore the intersection of the two concentric ellipsoids (3) *and* (4).

In a corresponding manner, we treat of the curves that the endpoints of the impulse vector and the rotation vector describe relative to fixed space. The former curve naturally reduces to a point; we therefore consider the latter; that is, our *herpolhode curve*. We first know from the theorem of the *vis viva* that this curve lies in the fixed plane e.

The herpolhode curve is therefore a plane curve. The distance of its plane from O (that is, the projection of the rotation vector on the vertical impulse axis), which we denote in agreement with the previous by ϱ, results from the theorem of the *vis viva* as

$$\varrho = \frac{2h}{G}. \tag{5}$$

The more precise form of the herpolhode curve may not, as it was for the polhode curve, be defined by an elementary geometric construction, since it is, in general, of a transcendental nature. But it is certainly possible to discover it from our kinematic image of the motion.

We describe spheres about O at the greatest and least distances of the polhode curve from O; these spheres determine two circles in our plane e. The common midpoint of these circles is the point of intersection of the impulse axis with e. Between these two circles the herpolhode curve must evidently run to and fro in regular windings, in which it is alternatively tangent to these circles or, in special cases, touches one of these circles with a cusp. The herpolhode curve consists of an infinite series of congruent arcs that are twisted with respect to one another by the same amount. Each individual arc corresponds to a single unwinding of the polhode curve. In general, the curve does not close, but rather encircles the midpoint of the figure, the intersection

point of the impulse axis with its plane, infinitely often. From this it already follows that the equation of the curve will be a transcendental equation. The corresponding result holds, naturally, for the herpolhode cone, which this curve projects from O. The form of the herpolhode is represented in the adjacent figure*) for the special case

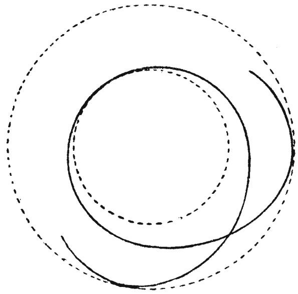


Fig. 16.

$$A = \frac{1}{36}, \quad B = \frac{1}{25}, \quad C = \frac{1}{16},$$

$$h = 50, \quad G^2 = 5.$$

We will now see how these matters are modified in the case of the *symmetric top.* Here the ellipsoid of inertia, as well as all the previously used ellipsoids (1), (3), and (4), has rotational symmetry about the figure axis. But if we bring together for intersection an ellipsoid of revolution and a sphere (as for the construction of the impulse curve), or two ellipsoids of revolution with coinciding figure axes (as for the construction of the polhode), then the intersection curve is always a pair of diametrically situated circles in parallel planes.

Thus the polhode curve, as well as the curve that the impulse describes with respect to the body, goes over into a circle.

If we let, further, an ellipsoid of revolution with fixed midpoint O roll on a plane, then the locus of the contact point in this plane is obviously a circle. This shows, for example, that the collected points of the polhode curve have a constant distance (Ω) from O; the collected points of the herpolhode curve must also lie at the distance Ω from O. The latter curve is therefore the intersection of a plane with a sphere of radius Ω.

Thus the herpolhode curve for a symmetric top also goes over into a circle.

The character of the resulting motion may now be given in a word:

The most general motion of the force-free symmetric top is regular precession.

---

*) The figure is taken from the dissertation of Mr. W. H e s s: "Das Rollen einer Fläche zweiten Grades auf einer invariabeln Ebene, München 1880." Mr. H e s s shows that the herpolhode curve, because of the inequalities that hold between the principal moments of inertia A, B, C, can have no (real) inflection points; the figure originally given by P o i n s o t, Liouville's Journal sér. I, t. 16 was erroneous in this respect.[66]

In fact, we could have characterized regular precession in §6 of the preceding chapter by the condition that the curves and cones of the polhode and the herpolhode were circles and circular cones, respectively. The axis of precession is the impulse axis. We will reserve for later the more precise classification of this precessional motion in the sense of the distinctions of page 53.

In the special case of the symmetric top, the following deliberation, which is perhaps even simpler and more direct, also leads to the goal. We again take the impulse axis vertically upward and draw the figure axis, in its initial position, inclined by an arbitrary angle ϑ with respect to the impulse axis. The position of the instantaneous axis of rotation follows from the positions of the impulse axis and the figure axis. The rotation axis always lies, according to the construction given on page 106, in the same plane as the impulse and figure axes, and divides the angle ϑ, as we can say concisely, in a fixed ratio depending only on the mass distribution of the top (that is, on the values of A and C), and, in particular, is interior or exterior according to whether the top is prolate or oblate. For greater clarity we lay a unit sphere about O and mark its points of intersection with the impulse axis, the rotation axis, and the figure axis, which we denote, respectively, by J, R, and F. The point J, according to our fundamental principle, is a fixed point, the "north pole" of the sphere. The points R and F, in contrast, move; *we claim that they each describe a parallel circle about the north pole.*

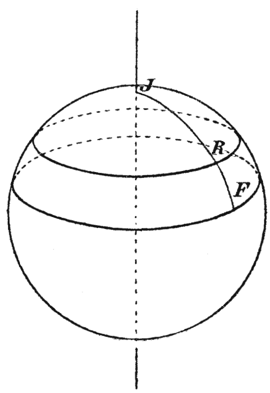


Fig. 17.

In fact (cf. Fig. 17), the instantaneous motion of the top consists of a rotation about OR. Since F and R lie on the same meridian of the sphere, the point F wanders at the first moment in the direction of the parallel circle through F, so that the angle ϑ will be initially unchanged. As a result, another line of the body that is contained in the meridian plane JOF must now take over the role of the turning axis. If we denote its intersection point with the unit sphere by R, then R lies on the meridian JF. Since this point now divides the arc JF in a fixed ratio and the arc JF must retain its initial length, the arc JR

must also have its original magnitude. Thus R likewise wanders, at the first moment, on the parallel circle through R, and indeed, as we said, so far that J, R, and F will be points on one and the same meridian. We are thus led back precisely to the initial conditions of the motion. In consequence, our deliberation also holds at each following time.

For force-free motion of the symmetric top, the figure axis and the rotation axis each describe a circular cone about the impulse axis.

From our earlier construction of the rotation vector, it follows, further, that the length of the rotation vector is likewise constant for constant length of the impulse vector and constant inclination angle ϑ. Now, however, the progressive velocity of the figure axis on its circular cone is proportional to the magnitude of the rotation vector. *The figure axis therefore traverses its circular cone with constant velocity.* Further, the rotational velocity of the top about the figure axis equals the projection of the rotation vector onto the figure axis. *The top therefore rotates relative to the figure axis with constant angular velocity.*

By these remarks, however, the motion is again characterized as *regular precession.* —

In conclusion, we remark that our treatment of the force-free top also finds application to the motion of a free rigid body in space that is subjected to no external forces or only such external forces that, by an appropriate choice of the reference point, may be composed into a pure pushing-force, as, for example, the force of gravity in the case where we choose the center of gravity as the reference point. Namely, we can then treat, according to the general impulse theorems of the fifth section, of the translation of the reference point on the one hand, and the rotation of the body about the reference point on the other hand, the latter according to the preceding theory of the force-free generalized top, and the former according to the laws of the mechanics of a single particle. Since for the case of the action of gravity, in particular, the trajectory of a single mass particle (the parabola) is sufficiently known, we now already command, with recourse to the preceding results, *the motion of a heavy rigid body moving freely in space.*

§8. Rotation of the top about a permanent turning axis and the so-called stability of the rotation axis of a rapidly rotating top

We have already emphasized many times the analogy between the motion of a single mass particle and the rotation of the top. Kinematically, both problems have three degrees of freedom; statically, the impulse in both cases can be conceived as a vector; a deep analogy also exists kinetically, in so far as we direct our attention to the behavior of the impulse vector (cf. §5). The situation is different if we compare the behavior of the velocity vector in the two cases. While the velocity vector (just as the impulse vector) retains its direction and magnitude in space for the single force-free mass particle, the vector of the turning velocity for force-free motion of the top continuously changes in magnitude and direction, both in space and in the body. We will pose the problem of determining the circumstances under which the velocity vector of the top also remains constant in magnitude and direction in space; or, in other words, *under what circumstances a uniform rotation of the top about a fixed spatial axis occurs.*

We know that for the *Poinsot motion* the rotation vector describes a conical shell that has the impulse vector in its interior. If the rotation axis is now stationary in space, the herpolhode cone reduces to a single line, so that this line must coincide with the direction of the impulse. According to page 101, however, the rotation axis and the impulse axis coincide only when their common direction is a principal axis of the body. Conversely, it follows from the construction of page 102 that the rotation axis then retains a fixed position in space and in the body, and that the rotational velocity is uniform. If we denote, as usual, an axis about which a continuous uniform rotation is possible as a "permanent axis," then we can say:

For the generalized top there are only three permanent axes, the principal axes of the body.

If the top rotates about one of these three principal axes, then the polhode, herpolhode, and the curve that the impulse describes in the body each reduce, obviously, to a single point.

The three principal axes present an interesting difference with

respect to the *stability* of the considered uniform rotation, as was already noted by P o i n s o t.

The concept of the stability of a motion, which we encounter here for the first time, plays an important role in modern mechanics, and will be treated with a certain care. To determine whether we call a form of motion of our top stable or unstable, we will proceed in the following way (where the precise sense of the words chosen by us will become fully clear, perhaps, only in the course of the further development): we impart to the top, while it executes the motion in question, a small impact of an arbitrary nature. If the collected elements of the motion (for example, the successive positions of the top in space and the positions of the rotation vector and the impulse vector with respect to the top and in space) are always changed to a lesser extent as the applied impact is made smaller, then we will call the motion stable; in every other case it will be called unstable.

In this spirit, we first investigate the *rotation of the top about the axis of the greatest or least principal moment of inertia*, and consider, above all, the *curve that the endpoint of the impulse describes in the body.* The ellipsoid 1) and the sphere 2) of page 123, whose common points yield our impulse curve, must, obviously, now be tangent at two points that lie on the longest or shortest principal axis of inertia, respectively. In the former case, the ellipsoid will be completely enclosed by the sphere, and in the latter the sphere will be completely enclosed by the ellipsoid. If we now impart a small impact to the top, then we slightly change the constants h and G; that is, the sizes of our ellipsoid and our sphere. The tangent point dissolves, in both cases, into a small encircling curve, which lies, in its entire extent, always closer to the previous tangent point as the changes in h and G are made smaller. (For arbitrary changes in h and G, the tangent point can dissolve, to be sure, into an imaginary curve; the necessary values of our constants, however, are incompatible with the mechanical meaning of these quantities, so that we can therefore disregard them.)

The *polhode* behaves in a manner entirely similar to the impulse curve; indeed, the polhode curve can be derived from the impulse curve by a simple deformation of the principal axes of the body. The polhode curve, which for uniform rotation consists of a single point, is also transformed by the addition of a small external impact into a small oval that continually lies very near the previous single point. We thus conclude that in the rolling of the polhode on our fixed plane of

page 121, there results a *herpolhode curve* whose dimensions are likewise always smaller as the disturbing impact is chosen smaller. In summary, it follows that:

Uniform rotation of the top about the greatest or least principal axis of the ellipsoid of inertia is a stable form of motion.

We next assume that the rotation occurs about the *intermediate principal axis of the ellipsoid of inertia.* Again, we first consider the impulse curve. This curve reduces, in our case, to one of the intersection points of the intermediate principal axis with the sphere of radius G. Our sphere 2) is tangent to the ellipsoid 1) at this point as well as at the diametrically opposed point. One sees immediately from the adjacent figure, however, that there are still infinitely many other points which are common to the two surfaces. Namely, the two surfaces must necessarily intersect; the ends of the greatest principal axes of the ellipsoid, for example, extend out of the sphere; the smallest principal axes lie entirely in the interior of the sphere. The complete intersection curve consists of two circles; namely, the intersection circles of the ellipsoid, which cross at the endpoints of the intermediate principal axis. It is easy to give the analytic condition that is required for the occurrence of the present case. If B is the moment of inertia corresponding to the intermediate principal axis, then we must choose the constants h and G so that equations (1) and (2) on page 123 give the same value of M^2 for $L = N = 0$; our condition is thus

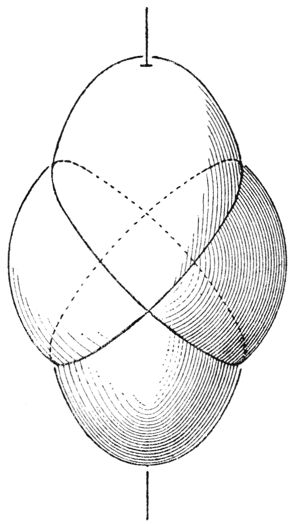


Fig. 18.

$$G^2 = 2hB.$$

If we now alter the dimensions of the sphere and ellipsoid slightly by the addition of an external impact, then we always obtain a curve that departs from the original tangent point by a finite distance. The two tangent points (together with the circles passing through them) dissolve into two ovals that surround the endpoints of the largest or smallest principal axis of the altered ellipsoid. This is seen in Figures 19 and 20. The external disturbance for these figures is specifically chosen

so that the ellipsoid retains its size and only the sphere is changed; in particular, the sphere is made larger with respect to the original in Figure 19 ($G^2 > 2hB$) and smaller in Figure 20 ($G^2 < 2hB$).

On the basis of the depicted behavior of the impulse curve, we can already state the interesting theorem:

Uniform rotation of the generalized top about the axis of intermediate principal moment of inertia is an unstable form of motion.

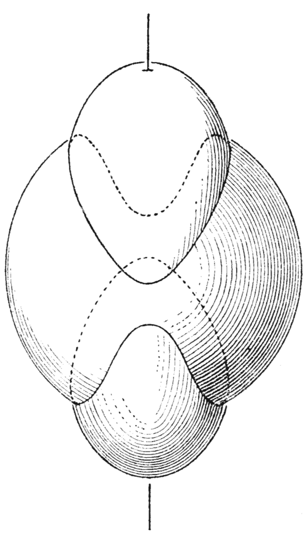


Fig. 19.

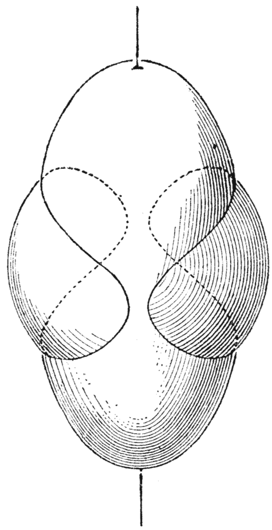


Fig. 20.

We also wish to consider briefly the forms of the polhode and herpolhode curves. The form of the polhode curve is, naturally, completely analogous to that of our impulse curve. In the case $G^2 = 2hB$, the polhode curve consists of a point, while the intersection of the ellipsoids 3) and 4) of page 124, on which the polhode runs, gives two congruent ellipses. (They result from the above intersection circles of the ellipsoid 1) by the deformation $p = L/A$, $q = M/B$, $r = N/C$.) If we now make $G^2 \gtrless 2hB$ by an external impact, the two ellipses dissolve into two ovals, which are removed from the point-shaped polhode of the case $G^2 = 2hB$ by a finite amount, however small the disturbing impact is.

From this behavior of the polhode curve, we immediately conclude that the herpolhode curve, which we have indeed obtained from the unrolling of the polhode curve, will also change its form discontinuously. While the herpolhode curve consists of a single point in the case $G^2 = 2hB$, it likewise attains, in the case $G^2 \gtrless 2hB$, a finite dimension

that cannot be made arbitrarily small by the diminishment of the difference $G^2 - 2hB$. We cannot, in this place, go more deeply into the interesting details*) that appear here.

We now go over to the *symmetric top. For the symmetric top, there are obviously infinitely many permanent rotation axes; these are, in addition to the figure axis, the collected axes of the equatorial plane.* We also pose here the question of the *stability* of the considered form of motion.

For *rotation about the figure axis*, the question is settled if we conceive the symmetric top as the limiting case of the generalized top. For the generalized top, the figure axis corresponds in every case to the axis of the largest or smallest principal moment of inertia (according to whether the symmetric top is prolate or oblate). Thus we can say:

Rotation about the figure axis is a stable form of motion of the symmetric top.

For *rotation about an axis of the equatorial plane*, in contrast, the comparison with the generalized top forsakes us. Such an axis, namely, can be regarded equally well as either the limiting case of the intermediate principal axis or the limiting case of one of the two extreme principal axes. Correspondingly, the following specific investigation shows that the stability properties of the symmetric top for this rotation hold, in a certain sense, the middle ground between the complete stability of a rotation about an axis of an extremal principal moment of inertia and the complete instability of a rotation about the axis of the intermediate principal moment of inertia.

Here, we consider first the *herpolhode curve.* We rely on our previous theorem that the most general motion of the symmetric top is a regular precession about the axis of the impulse. To begin with, the impulse lies exactly on an equatorial axis. The herpolhode curve thus consists of a point that lies on this same axis. We then add a small additional impulse, by which the position of the impulse in space and in the body will be slightly altered instantaneously. The position of the rotation vector follows from our familiar construction. In every case, the altered vector is very slightly displaced in magnitude and direction from the

*) Under some circumstances, the herpolhode curve assumes a spiral form. This occurs if the impact is specifically chosen so that the previously mentioned intersection curve of the ellipsoids 3) and 4) (consisting of our two congruent ellipses) is not changed as such, and the rotation pole in the body is only displaced slightly from the double point of the intersection curve to one or the other of the ellipses.[67]

original. We now obtain the form of the herpolhode curve if we lead the endpoint of the thus constructed rotation vector in a small circle about the changed axis of the impulse. This circle always becomes smaller for a smaller disturbing impact. The herpolhode curve therefore changes its form arbitrarily little – in contrast to the herpolhode curve for the corresponding motion of the generalized top.

Nevertheless, however, we must designate the motion in question for the symmetric top as unstable according to our above definition of stability. Namely, we consider the *polhode curve.* For the original position of the impulse, the polhode curve consists of a single point. This statement appears to be in contradiction to the earlier result that the polhode curve is always a circle about the figure axis. The contradiction is resolved if we consider that in uniform rotation about an equatorial axis the endpoint of the rotation vector does have the tendency to progress in a circle about the figure axis, but does not leave its place here because the herpolhode curve consists of a single point. We now obtain the altered form of the polhode curve due to our impact if we let the altered endpoint of the rotation vector move in a circle about the figure axis. With the addition of our small impact, the polhode curve thus changes its form in a discontinuous way — in agreement with the polhode curve for the corresponding motion of the generalized top. The same holds for the curve that the impulse describes with respect to the body.

According to the preceding, we must say, in any case:

Uniform rotation of the top about an axis in its equatorial plane is an unstable form of motion.

Finally, we consider the *spherical top.* Here each axis is a principal axis of the ellipsoid of inertia, from which it follows:

For the spherical top, each axis through O is a permanent rotation axis.

The most general motion of the spherical top consists of a uniform rotation about an axis fixed in space.

One is easily convinced, moreover, that each such rotational motion has a *stable* character.

If we take up once again the comparison between the kinetics of the top and that of a single mass particle posed at the beginning of this section, then we can say:

The spherical top forms, not only with respect to the behavior of the impulse vector but also with respect to the velocity vector, an exact analogue to the single mass particle; namely, the magnitude and direction of this vector also remain constant in space for force-free motion of the spherical top. —

The word "stability" is frequently used, according to the precedent of Foucault, in a sense different from that given above. The word is used, namely, to mean the apparent tendency of a uniformly rotating body to retain the spatial direction of its rotation axis in the presence of external disturbances. The experiments that concern this phenomenon are sufficiently well known. We imagine, for example, a top such as our demonstration model depicted in the Introduction, which we may balance by upper weights so that the center of gravity lies at the support point. If we set the top in motion by unwinding a string, we impart to it a rather strong rotation that has the figure axis as the rotation axis. Without great exertion, it is attained that the top makes 20 revolutions per second. If we now wish to change the inclination of the figure axis discernibly, we must apply a considerable force. Smaller disturbances, such as a shaking of the pedestal or a light blow to the upper surface of the top, hardly cause a noticeable change in the state of motion. We compare this to the fact that the nonrotating top obviously reacts with an evident motion to any arbitrary impact. We will then, in fact, be inclined to believe in a certain capacity for resistance that the top acquires by virtue of its rotation.

Analogous phenomena are observed very frequently for freely moving bodies*). If one would strike a target with a thrown body, then one always hurls the body with as strong a rotation as possible. Through this alone can one produce a regular and predictable trajectory of the body. In other cases, all sorts of accidental circumstances, such as the effect of air resistance, incidental currents in the air, etc., would disturb the path considerably. This applies, for example, to the discus throwing of ancient times or the hoop game common today. The effect finds great employment in the construction of modern artillery and

*) Numerous examples of this type are found in the popularly received work of Perry: Spinning tops, London 1890; we wish to earnestly recommend this interesting and, through its amusing presentation, distinguished little book.[68]

infantry weapons; namely, in the use of *rifled* barrels, to which we will later return.

The explanation of all these phenomena is, for us who command the concept of the impulse, exceedingly simple. We speak in this respect of our top, but we could just as well think of any one of the given examples. Through the original impetus, we have created an impulse vector that lies nearly in the direction of the figure axis and has a considerable length. This impulse is composed with the impulse of the external disturbance according to the parallelogram of forces. If the latter is considerably smaller, as in the previous examples, then the original impulse will be changed only slightly in magnitude and direction. Thus the changed state of motion also differs only slightly from the original; the rotation axis retains its position in space approximately, and the figure axis also remains in the proximity of its earlier position.

To give a numerical example, we may take the previously cited number of 20 rotations in one second as a basis. The angular velocity then amounts to $2\pi\,.\,20$ (sec^{-1}). It corresponds, according to our convention, to a rotation vector of length $2\pi\,.\,20$ cm. To calculate the corresponding impulse vector, we must make a judgment of the magnitude of the moment of inertia about the figure axis. We imagine, for this purpose, the bell-shaped principal piece of our model to be replaced, for example, by a circular disk of 1 cm thickness and 10 cm radius. The density of the material is approximately 8 (gr. cm^{-3}), (namely, 7,6 for iron, 8,5 for copper). For the moment of inertia, one thus easily finds the value $4\pi\,.\,10^4$ (gr. cm^2); the impulse vector thus receives, according to our earlier agreement, the length of $(4\pi)^2 10^5$, approximately equal to $15\,.\,10^6$ cm. The impulse vector is thus approximately 150 kilometers or 20 German miles long! We next calculate the magnitude of the additional impulse in a concrete example. We wish, for example, to let a body with mass one gram fall onto the edge of the circular disk of the top from a height of one-half meter. The velocity that our body attains is equal, according to the falling law, to $\sqrt{2gh}$; that is, for $h = 50$ cm and $g =$ ca. 900 (cm sec^{-2}), equal to 300. The applied impact thus amounts to 300 (gr. cm sec^{-1}). The turning moment itself, since the radius of the circular disk was taken to be 10 cm, is equal to 3000. Our additional impulse thus has a length of 30 m and is, if the figure axis stands vertical, directed horizontally. It is clear that this additional impulse, compared with the imposing length

of the original impulse, signifies only a relatively insignificant change. The corresponding change in the state of motion will thus be hardly observable. In summary, we can say:

The phenomena in question are, in themselves, no more remarkable than the thoroughly self-evident fact that a rapidly moving body will be deflected in direction by a lateral impact ever more slightly as the original translational impulse, or, equivalently, its translational velocity, becomes larger. —

The situation may appear somewhat differently if we have, instead of a single or short disturbance, a continuous influence. Obviously, a quite small but persisting change in impulse can produce a considerable effect even compared to a very large initial value, if only we extend the observation over a sufficiently large time interval. This is in fact the case, as is shown in the sequel for the example of the gravitational influence.

After we have completely understood the phenomena themselves, we will, finally, criticize the expression *stability of the rotation axis* by which one frequently describes these phenomena.

First, we wish to reserve for the word "stability" the special sense that was given at the beginning of this section, and that forms the opposite of the word "lability." We will thus prefer to say, instead of stability of the rotation axis, *"conservation of the rotation axis."* Next, we prefer, in general, not to speak here of the rotation axis, but rather of the impulse axis. *It is not the rotation vector, but rather the impulse vector that takes the first place dynamically.* If we have determined the position of the impulse vector, the position of the rotation vector follows immediately, for example, on the basis of the construction of page 101. It is entirely false that an external cause, a force, is necessary for a displacement of the rotation vector. *The force is used not for the displacement of the rotation vector, but rather for the displacement of the impulse vector.* We think, for example, of the general motion for the force-free top. Here the rotation axis progressively changes its direction, in that it wanders along the herpolhode cone without the action of an external force, while, on the other hand, the impulse vector remains fixed. The change of the rotation axis occurs just because, and in just such a manner that, the impulse axis can remain unchanged.

And now (as we know), the "law of the conservation of the impulse axis" against external disturbances is not exactly valid. However small

the added impulse compared to the original may be, it always produces a finite change. This is enough to separate the perhaps initially coinciding axes of the rotation and the impulse. In consequence, the rotation axis is no longer stationary in space, but rather describes (for the symmetric top) a thin circular cone about the impulse axis. The same holds for the motion of the figure axis.

We will therefore be obliged to correct the alleged law of the stability of the rotation axis; we say:

The impulse will, under certain conditions frequently present in practice, be changed relatively little by external disturbances. As a consequence, if the motion is initially about the figure axis, the herpolhode cone and the cone described by the figure axis in the altered motion retain a very small opening after the disturbance.

Chapter III

The Euler equations, with further development of the kinetics of the top

§1. Derivation of the Euler equations

After having become acquainted in the previous chapters with the geometric and mechanical foundations of the theory of the top, we wish to begin this chapter by preparing the analytic treatment of the problem. It is necessary, in the first place, to derive the famous *Euler equations.*

We start with a kinematic consideration that is tied to our earlier results on infinitely small rotations.

We consider an arbitrary vector, which we refer, on the one hand, to a fixed coordinate frame at O, and, on the other hand, to a moving coordinate frame with the same origin. In the former case, we call the vector i, and in the latter case, J. Let the coordinates of its endpoint be l, m, n and L, M, N, respectively. The notation is chosen with the consideration that we will soon identify our vector with the impulse vector. The direction and magnitude of the vector may be arbitrarily variable. The change of the vector in the time dt will be denoted by di or dJ, respectively, according to whether we measure the change in the fixed or the moving frame. The components of di in the fixed coordinate frame are dl, dm, dn, and the components of dJ in the moving frame are dL, dM, dN. Finally, we denote by $d\lambda, d\mu, d\nu$ the components of di with respect to the moving coordinate axes. The moving coordinate frame experiences, at each moment, an infinitely small rotation R, whose components in the moving frame will be denoted, as previously, by p, q, r.

We now draw upon the schema (3) of page 41. We can interpret this schema as giving the relation between the coordinates of a fixed spatial point P with respect to two successive positions of the moving coordinate frame. As the point P, we consider the endpoint of the

R.J. Nagem, G. Sandri, *The Theory of the Top*, DOI: 10.1007/978-0-8176-4721-6_III,

vector i at time t. This point has coordinates L, M, N at time t with respect to the moving frame; we ask for the coordinates of the same spatial point at the time $t + dt$.

The desired coordinates are not simply $L + dL$, $M + dM$, $N + dN$. These coordinates, rather, correspond to the endpoint P' of the vector i, which has changed during the time dt. But we return from the point P' to P if we make the change of our vector di again in reverse. Now since the components of di in the moving frame are denoted by $d\lambda, d\mu, d\nu$, the coordinates of P with respect to the moving frame at time $t + dt$ will be

$$L + dL - d\lambda, \quad M + dM - d\mu, \quad N + dN - d\nu.$$

Our schema (3) now yields, read from above to below, the following relations for the coordinates of point P:

$$\begin{aligned} L + dL - d\lambda &= L + r\,M\,dt - q\,N\,dt, \\ M + dM - d\mu &= -\,r\,L\,dt + M + p\,N\,dt, \\ N + dN - d\nu &= q\,L\,dt - p\,M\,dt + N. \end{aligned}$$

Thus we can also write

$$(1) \qquad \begin{cases} \dfrac{dL}{dt} - \dfrac{d\lambda}{dt} = + rM - qN, \\ \dfrac{dM}{dt} - \dfrac{d\mu}{dt} = -\,rL + pN, \\ \dfrac{dN}{dt} - \dfrac{d\nu}{dt} = qL - pM. \end{cases}$$

The meaning of these equations becomes much clearer if we avail ourselves of the notation of vector analysis. If we use, namely, the concept of the vector product explained on page 61, then we can combine the previous equations in the form

$$(1') \qquad \frac{dJ}{dt} - \frac{di}{dt} = V(J, R),$$

where we may, naturally, also write $V(i, R)$ on the right-hand side instead of $V(J, R)$.[69]

The geometric difference of the rates of change of our vector with respect to the moving and the fixed frames is therefore equal in magnitude and direction to the vector product of our vector J and the rotation vector R.

If we combine this fundamental and very general kinematic theorem with the kinetic principles of the previous chapter, then the Euler equations result with one blow.

Force-free motion is considered first. Let the vector of the previous theorem be the impulse vector of the top; the coordinate frame XYZ may, at first, have an arbitrary position with respect to the top.

For force-free motion, the impulse vector is constant in space. We thus have

$$i = \text{const.} \quad \text{or} \quad \frac{di}{dt} = 0.$$

Equation (1) now gives

$$\frac{dJ}{dt} = V(J, R). \tag{2}$$

We have thus derived *a first vectorial form of the Euler equations for the force-free top.*

We obtain a second form of the same equations if we return from the vector equation (2) to its component equations with respect to the coordinate axes. It is this same transition, in the reverse sense, that led us above from (1) to (1′). Our *second form of the Euler equations for the force-free top* is thus

$$\left\{\begin{aligned} \frac{dL}{dt} &= \qquad\quad rM - qN, \\ \frac{dM}{dt} &= -rL \qquad\quad + pN, \\ \frac{dN}{dt} &= \quad qL - pM. \end{aligned}\right. \tag{2′}$$

We arrive, finally, at a third form of the same equations if we eliminate one of the two vectors J and R by virtue of the relation existing between them. The elimination is particularly simple if we let the XYZ frame coincide with the principal inertial frame. Namely, we can then replace the components of J simply by the values

$$L = Ap, \quad M = Bq, \quad N = Cr,$$

and obtain for the components of R the differential equations

$$\left\{\begin{aligned} A\frac{dp}{dt} &= (B - C)qr, \\ B\frac{dq}{dt} &= (C - A)rp, \\ C\frac{dr}{dt} &= (A - B)pq. \end{aligned}\right. \tag{2″}$$

These are the *Euler equations for force-free motion in the commonly used form.*

The latter form of the Euler equations is evidently less general than the preceding, since it assumes a special position of the coordinate frame. Equations (2′) are valid for an entirely arbitrary position of the frame fixed in the top, and do not change their form in the least if we rotate the coordinate frame. Equation (2) is, in a certain sense, the most perfect expression of the Euler equations, since it contains absolutely no reference to a coordinate frame. This equation immediately establishes the following fact:

The Euler equations for force-free motion are nothing other than the analytic statement that the impulse is constant in space.

The Euler equations in the presence of arbitrary external forces follow in the same way. We first compose the external forces into a single turning-force $\mathsf{\Delta}$ with respect to the reference point O. The turning-force $\mathsf{\Delta}$ has components $\mathsf{\Lambda}$, M, N in the (arbitrarily situated) XYZ frame. The spatial change di of the impulse due to the presence of the external forces in the infinitely small time dt is equal, according to the fundamental theorem IIa of page 115, to the infinitely small impact $\mathsf{\Delta}\,dt$. We thus have the equation

$$\frac{di}{dt} = \mathsf{\Delta}.$$

The Euler equations in the presence of external forces now follow directly from the kinematic equation (1′).

As *a first form of these equations* we obtain

$$\frac{dJ}{dt} = V(J, R) + \mathsf{\Delta}. \tag{3}$$

If we go over, as above, to the component equations, then there results a *second form of these equations*, which holds for an arbitrary position of the coordinate frame XYZ; namely,

$$\left\{\begin{aligned} \frac{dL}{dt} &= \quad rM - qN + \mathsf{\Lambda}, \\ \frac{dM}{dt} &= -rL + pN + \mathsf{M}, \\ \frac{dN}{dt} &= qL - pM + \mathsf{N}. \end{aligned}\right. \tag{3′}$$

If we wish to eliminate one of the two vectors J and R, then it is again convenient to let the XYZ frame coincide with the principal inertial frame. We then obtain the *usual form of the Euler equations in the presence of external forces:*

$$(3'') \qquad \begin{cases} A\dfrac{dp}{dt} = (B-C)qr + \mathsf{\Lambda}, \\ B\dfrac{dq}{dt} = (C-A)rp + \mathsf{M}, \\ C\dfrac{dr}{dt} = (A-B)pq + \mathsf{N}. \end{cases}$$

The forms (3) and (3′) of the Euler equations are again preferable to (3″), since they are independent of the choice of the coordinate frame. In particular, equations (3) show clearly that

In the presence of external forces, the Euler equations are the analytic expression of the fact that the rate of change of the impulse in space is equal in magnitude and direction to the turning-force corresponding to the external forces.

The equations named after Euler appeared for the first time in his treatise "Mouvement de rotation des corps solides autour d'un point fixe"*). The vector notation is due to Mr. Tait in his work**) "On the rotation of a body about a fixed point." The conception of the Euler equations adopted here was first (1856) given by Hayward***) in the previously cited treatise. It operates with the mechanical system in its entirety, and leaves nothing to be desired in simplicity. The Poinsot derivation of the same equations†), which we will criticize in detail directly, is much less transparent.

We wish to reproduce the corresponding Poinsot consideration only in so far as it provides us with a *new interpretation of the vector product of the impulse vector and the rotation vector that appears in the Euler equations.* We wish to show that this product, *kinematically* defined in our derivation, has a simple *kinetic* meaning.

For this purpose, we suppose at once that the rotation axis is fixed in the body and that the top rotates about this axis with constant velocity. Each mass particle of the top will then be subject to a centrifugal force whose magnitude and direction are determined according to the laws

*) Abhandlungen der Berliner Akademie vom Jahre 1758.[70]

**) Cf. Transactions of the R. Soc. of Edinburgh, Vol. 25, 1869, pp. 279, 280.[71] The vector $\varphi(R)$, which for Tait as well as earlier for Hamilton (Elements of Quaternion Calculus, Vol. 2, p. 350 of the German edition) is defined through the rotation vector by a rather complicated vector equation, is our impulse.

***) Cf. above, p. 113.

†) Cf. Théorie nouvelle ..., Nr. 53 and ff.

of elementary particle mechanics. We will compose all these centrifugal forces, with respect to the fixed point O as the reference point, into a resulting *"centrifugal turning-force"* C. *This turning-force,* we claim, *is directly equal to the vector product in question:*

$$C = V(J, R). \tag{4}$$

As a proof, we will proceed as on page 94, where we composed the total impulse of the top from the system of the impulses corresponding to the individual mass particles. Namely, we wish to take the (thought to be fixed) rotation axis as the X-axis, so that we can denote the (assumed to be constant) rotational velocity by p. Each mass element dm of the top will then travel in a circle about the X-axis with the constant angular velocity p. The radius of the circle, if the coordinates of dm are XYZ, will be $\sqrt{Y^2+Z^2}$.

The individual mass elements are subject, in this motion, to a centrifugal force whose vector will be denoted by K. We will develop the concept of the centrifugal force in the fifth section from the standpoint of the impulse theory. Here, we take from that development the well-known fact that the vector of the centrifugal force is oriented in direction and sense radially outward from the midpoint of the circle, and that the magnitude in our case is

$$|K| = dm \cdot p^2\sqrt{Y^2+Z^2}\ ^{*}).$$

The components of K with respect to the coordinate axes are then, as one easily recognizes,

$$K^X = 0, \quad K^Y = p^2 Y\, dm, \quad K^Z = p^2 Z\, dm.$$

The centrifugal turning-force C is computed from the system of individual forces K according to the laws of statics. From equations (1) on page 85, the components of C are given by

$$C^X = 0, \quad C^Y = -p^2 \int ZX\, dm, \quad C^Z = p^2 \int XY\, dm. \tag{5}$$

*) We take the symbol $|K|$ for the magnitude of the vector K from the usual notation of function theory (thanks to which one denotes the absolute value of the complex number $z = x + iy$, that is, the length of the two-term vector $x + iy$, by $|z|$). The introduction of this symbol will frequently be convenient in the following, since it is tiresome to express the direction and magnitude of the vectors in our formulas in each case.

We recall, on the other hand, the expressions that were derived on page 94 for the impulse that results for the same position of the coordinate frame. We found for the components of this impulse

$$(5') \quad L = p\int (Y^2+Z^2)\,dm, \quad M = -p\int XY\,dm, \quad N = -p\int ZX\,dm.$$

With consideration of these equations, we can write equations (5) as

$$C^X = 0, \quad C^Y = pN, \quad C^Z = -pM.$$

But the quantities on the right-hand sides are, according to the analytic definition of the vector product on page 62, directly identical to the components of $V(J, R)$, since, for our choice of the coordinate frame, two components of the vector R vanish ($q = r = 0$). Our claim above is thus proven.

On the basis of the previously mentioned geometric meaning of the vector product, we can thus state the following theorem, which naturally remains valid if we abandon the special choice of the coordinate frame that was made, for convenience, above:

The resultant centrifugal turning-force is represented by a vector that is simultaneously perpendicular to the rotation vector and the impulse vector, and lies on that side from which the latter is transformed into the former by a positive rotation along the shortest path. The magnitude of our turning-force in the absolute measurement system is equal to the area of the parallelogram constructed from the impulse vector and the rotation vector.

A *new interpretation of the Euler equations*, finally, is associated with this kinetic conception of the term $V(J, R)$. But we emphasize explicitly that we ascribe only a secondary interest to this new interpretation compared to our original, according to which the Euler equations state the invariance of the impulse in space.

We first remark that we can write, with consideration of (4), the Euler equations for force-free motion, or for motion influenced by the external turning-force Δ, as

$$\frac{dJ}{dt} = C \quad \text{or} \quad \frac{dJ}{dt} = C + \Delta.$$

Thus we can state, as addenda to the impulse theorems Ia and IIa of page 112 and page 115, two new theorems that refer to the change of the impulse relative to the top:

T h e o r e m I b. *For force-free motion, the impulse is changed relative to the body by the composition, at each moment, with the infinitely small turning-impact produced by the centrifugal forces just considered.*

T h e o r e m I I b. *For motion under the influence of external forces, the change in the impulse relative to the body is equal, at each moment, to the turning-impact produced by our centrifugal forces augmented by the turning-impact corresponding to the external forces.*

These impulse theorems show that one can treat of the inverse motion, which is determined by the impulse J and its change $\frac{dJ}{dt}$, in the same way as the direct motion, if only one adds the external forces of the direct motion in question to the centrifugal forces that would act on the individual mass elements in an assumed constant rotation.

The content of our impulse theorems becomes particularly plausible in the special case $\frac{dJ}{dt} = 0$, *where the impulse is fixed in the body.* The corresponding motion obviously consists of a uniform rotation about a fixed (in the body and in space) axis. Here it is completely clear that we must — in conformity with impulse Theorem IIb — apply from the exterior a turning-force $\Delta = -C$, which is equal and opposite to the centrifugal turning-force.

The new interpretation of the Euler equations is now simply that we reverse the order of the latter considerations and say that *the Euler equations are the direct analytic expression of our impulse Theorems Ib and IIb.*

For P o i n s o t, this latter interpretation of the Euler equations stands in the foreground. His treatment of the force-free motion of the top rests directly on the axiomatically presented, and not completely clearly formulated, impulse Theorem Ib. But this impulse theorem is obviously much less easily understandable than our earlier theorems that referred to the change of the impulse in space. One notes, above all, that the centrifugal forces of which our current theorem treats come into effect only if one imagines the rotation axis to be fixed in the body, and that they do not appear at all, therefore, in the actual motion of the top (for which the rotation axis changes continuously). It therefore seems to us improper to state, with Poinsot, these new impulse theorems as self-evident axioms, and to base the derivation of the Euler

equations on them. Correspondingly, most textbooks*) also abandon the Poinsot derivation of the Euler equations, and give a treatment that is identical in spirit with our original derivation, and appears slightly different only because the meaning of the impulse concept is not presented, in principle, with sufficient clarity from the beginning.

In conclusion, a place may be found for a few further remarks that are associated with the concept of the centrifugal turning-force C.

We first consider the justification of the words "centrifugal moment" that we assigned on page 98 to the integrals

$$\int YZ\,dm, \quad \int ZX\,dm, \quad \int XY\,dm.$$

The calculation of C given above shows that these quantities are directly, or at least up to the sign, equal to the components of the turning-force arising from the individual centrifugal forces if we turn the body with angular velocity 1 about one of the coordinate axes. *The quantities designated as centrifugal moments therefore actually appear, under certain conditions, as turning-moments of the centrifugal forces.*

If the rotation axis coinciding with the X-axis is, for example, a principal axis of the body, then, according to the definition of the principal axes, the two centrifugal moments

$$\int ZX\,dm \quad \text{and} \quad \int XY\,dm$$

are equal to zero; as a result, according to (5), all three components of C vanish. *The masses of the body are balanced about the principal axes, so that the centrifugal forces of the individual mass elements cancel.* In this case, there is no reason, for force-free motion, that the impulse vector should change its position with respect to the body. We thus arrive at the familiar fact that the principal axes are permanent impulse axes, and therefore permanent rotation axes.

In all other cases, the resulting centrifugal turning-force has a nonzero magnitude and is perpendicular to the impulse axis and the rotation axis. According to our impulse Theorem Ib, it therefore causes during the time dt, in the case of force-free motion, a repositioning of the impulse vector relative to the body, in which the endpoint of the impulse vector will be displaced by an increment $C\,dt$. *The endpoint of the impulse thus progresses, with respect to the body, in a direction*

*) Cf., for example, Despeyrous-Darboux, Appell, etc.

that is always simultaneously perpendicular to the instantaneous rotation vector and the impulse axis, with a velocity that is equal to the parallelogram of the two vectors. —

Finally, a last important remark about the Euler equations. The Euler equations determine the position of the impulse vector (or that of the rotation vector) with respect to the top. To determine, on the other hand, the position of the top in space, a further system of equations is required, which is of a purely kinematic nature, and does not depend on the nature of the external forces acting on the top. According to whether the position of the top is written in terms of the parameters α, β, γ, δ or φ, ψ, ϑ, we will use, for this purpose, equations (4) of page 43 or equations (9) of page 45. *The complete differential equations of the top problem are thus represented by the Euler equations on the one hand, and by one of the cited systems of equations on the other hand.*

§2. Analytic treatment of the force-free motion of the top

We now wish to derive once again and complete the theory of the force-free motion of the top, which has already been discussed in the previous chapter. We begin from the differential equations of the problem; that is, from the Euler equations, on the one hand, and the kinematic equations written in terms of φ, ψ, ϑ, on the other hand. We thus have first

$$(1)\qquad \begin{cases} \dfrac{dL}{dt} = Mr - Nq, \\[2mm] \dfrac{dM}{dt} = Np - Lr, \\[2mm] \dfrac{dN}{dt} = Lq - Mp; \end{cases}$$

but further

$$(2)\qquad \begin{cases} \dfrac{d\varphi}{dt} = r - \dfrac{\cos\vartheta}{\sin\vartheta}\,(p\sin\varphi + q\cos\varphi), \\[2mm] \dfrac{d\psi}{dt} = \dfrac{1}{\sin\vartheta}\,(p\sin\varphi + q\cos\varphi), \\[2mm] \dfrac{d\vartheta}{dt} = (p\cos\varphi - q\sin\varphi). \end{cases}$$

The matter now is to integrate these equations. We remark in advance that the integration process is divided into two steps. Namely, we can treat of equations (1) in themselves, and from them determine p, q, r

as functions of time. With the expressions thus found, we enter the differential equations (2), which yield the values of φ, ψ, ϑ.

The integration of the Euler equations (1) consists in finding three independent relations that contain t and p, q, r, but not their differential quotients with respect to time. Two such relations can be established immediately in algebraic form.

If we multiply equations (1) by p, q, r or by L, M, N, respectively, and add, then the right-hand side becomes zero both times. Thus we have

$$\left\{\begin{aligned} p\,dL + q\,dM + r\,dN &= 0, \\ L\,dL + M\,dM + N\,dN &= 0. \end{aligned}\right. \tag{3}$$

The meaning of these equations is well known to us. They correspond to the fact, emphasized on page 146, that direction of progression of the endpoint of the impulse in the body is always simultaneously perpendicular to the rotation vector and the impulse.

Equations (3) can now, with consideration of the relation between the impulse vector and the rotation vector, be integrated immediately; written in terms of p, q, r, the result of the integration is

$$\left\{\begin{aligned} Ap^2 + Bq^2 + Cr^2 &= 2h, \\ A^2p^2 + B^2q^2 + C^2r^2 &= G^2, \end{aligned}\right. \tag{4}$$

in agreement with page 124.

The *polhode curve* is defined by equations (4) as a *space curve of the fourth order.* If we construct from the given equations a homogeneous relation in p, q, r, then this equation yields the *polhode cone.* For this purpose, we multiply the first equation by G^2, the second by $2h$, and obtain by subtraction

$$(AG^2 - 2hA^2)p^2 + (BG^2 - 2hB^2)q^2 + (CG^2 - 2hC^2)r^2 = 0.$$

The polhode cone will therefore be, in our case, a *cone of the second degree.*

The complete integration still lacks a third equation between p, q, r, and t. For the sake of symmetry, we wish to establish this third relation in such a way that we express the dependence of the instantaneous rotational velocity Ω on t. This dependence, however, is not of an algebraic nature, but rather leads, as we will see immediately, to elliptic transcendentals. We thus find it understandable that our earlier elementary geometric treatment regarding the temporal development of

the Poinsot motion must remain incomplete, and that the earlier treatment is not fully sufficient, in particular, for the representation of the angular velocity at each instant.

We proceed in the simplest way as follows. We combine equations (4) with the identity

$$\Omega^2 = p^2 + q^2 + r^2 = u, \tag{4'}$$

in which u is an auxiliary variable. From (4) and (4′), p^2, q^2, and r^2 are calculated as linear functions of u, and indeed one finds, by a small determinant calculation,

$$\begin{cases} p^2 = \dfrac{uBC - 2h(B+C) + G^2}{(A-B)(A-C)}, \\[2mm] q^2 = \dfrac{uCA - 2h(C+A) + G^2}{(B-C)(B-A)}, \\[2mm] r^2 = \dfrac{uAB - 2h(A+B) + G^2}{(C-A)(C-B)}. \end{cases} \tag{5}$$

In particular, the product pqr will therefore be equal to the square root of an expression of the third degree in u.

We now return to equations (1). We multiply these equations by $\frac{2p}{A}$, $\frac{2q}{B}$, $\frac{2r}{C}$, and obtain by addition

$$2\Big(p\frac{dp}{dt} + q\frac{dq}{dt} + r\frac{dr}{dt}\Big) = \frac{du}{dt} = cpqr,$$

where we set, as an abbreviation,

$$c = 2\Big(\frac{B-C}{A} + \frac{C-A}{B} + \frac{A-B}{C}\Big).$$

There follows

$$dt = \frac{1}{c}\frac{du}{pqr} \quad \text{or} \quad t - t_0 = \frac{1}{c}\int_{u_0}^{u} \frac{du}{pqr}. \tag{6}$$

We must designate the integral appearing here, according to what was just said concerning the dependence of the product pqr on u, as an *elliptic integral*, in so far as the denominator represents the square root of an expression of the third degree in u. We wish, however, to go more deeply into the theory of elliptic integrals only in the following chapter on the problem of the heavy symmetric top. We merely remark here that t is determined as a function of u by equation (6), and therefore also, in reverse, u or Ω^2 as a function of t. By means of equations (5), we can thus calculate p, q, r as functions of t.

We now consider equations (2), in which we regard p, q, r as known. We can circumvent the direct integration of these equations if we rely

on the fact that the impulse is constant in space. We can choose the fixed axis of the impulse as the z-axis, so that the endpoint of the impulse receives the coordinates $x = 0$, $y = 0$, $z = G$ and $X = Ap$, $Y = Bq$, $Z = Cr$. The direction cosines of the z-axis with respect to the axes of X, Y, Z, which were previously denoted by c, c', c'', are therefore

$$c = \frac{Ap}{G}, \quad c' = \frac{Bq}{G}, \quad c'' = \frac{Cr}{G};$$

on the other hand, we have, according to equations (6) on page 20,

$$c = \sin\vartheta \sin\varphi, \quad c' = \sin\vartheta\cos\varphi, \quad c'' = \cos\vartheta.$$

Thus

$$\sin\vartheta\sin\varphi = \frac{Ap}{G}, \quad \sin\vartheta\cos\varphi = \frac{Bq}{G}, \quad \cos\vartheta = \frac{Cr}{G}. \tag{7}$$

These equations determine φ and ϑ for each value of t if, as we assume, p, q, r are known functions of time. They must naturally be compatible with equations (2), and represent two integrals of these equations.

Concerning the associated arbitrary constants of integration, we have implicitly disposed of them when we took the impulse axis specifically along the axis of z.

The angle ψ is then calculated by a mere quadrature. Namely, the middle of equations (2) yields, with consideration of (7),

$$\frac{d\psi}{dt} = G\,\frac{Ap^2 + Bq^2}{G^2 - C^2r^2},$$

which we can also write, according to equation (4), as

$$\frac{d\psi}{dt} = G\,\frac{Ap^2 + Bq^2}{A^2p^2 + B^2q^2} = G\,\frac{2h - Cr^2}{G^2 - C^2r^2}. \tag{8}$$

Since p, q, r are regarded as known functions of t, ψ can thus be found for any value of t by a quadrature.

Through the preceding formulas, often developed in textbooks, it is possible to numerically calculate the complete elements of the force-free motion of the top. We could now, for example, answer the previously posed question regarding the form of the herpolhode curve in detail, if only we did not have to defer, for the present, the study of elliptic integrals.

Moreover, it would be advantageous to base an extensive analytic treatment on our parameters α, β, γ, δ instead of φ, ψ, ϑ. It would thus be shown, as we will later describe for the theory of the heavy symmetric top, that the formulas in terms of α, β, γ, δ become not only more symmetric, but also considerably simpler in function-theoretic

respects, than those in the commonly used φ, ψ, ϑ. One notes that the kinematic equations written in terms of α, β, γ, δ already have an advantage over the corresponding equations in φ, ψ, ϑ. While the equations for φ, ψ, ϑ represent an irreducible system of three differential equations of the first order, the equations for α, β, γ, δ are divided into two pairs of two simultaneous linear differential equations, which allow the determination of the quantities α and β on the one hand, and the quantities γ and δ on the other.[72]

Continuing in this way, we would also obtain the beautiful results of Jacobi*), which directly represent the nine direction cosines a, b, $c, \ldots$ in their dependence on t as ϑ-quotients, and would discover a new illumination of the associated important investigation of Hermite*).

We now wish to investigate the very simple, as we know, *force-free motion of the symmetric top* from the differential equations. We begin again with equations (1) and (2), in which we set $A = B$. There then follow from the integrals (4), through appropriate combinations,

$$(4') \qquad \begin{cases} p^2 + q^2 = \text{const.} = \dfrac{2hC - G^2}{A(C-A)}, \\[2ex] \quad r^2 \quad\; = \text{const.} = \dfrac{2hA - G^2}{C(A-C)}. \end{cases}$$

These equations show that the polhode curve now consists, in a well-known manner, of a circle described about the figure axis. At the same time, we see that the quantity

$$u = p^2 + q^2 + r^2$$

now becomes, in its turn, constant. As a result, the above method for finding a third integral, in which one integrates with respect to the variable u, is no longer applicable. But we very easily find the necessary third integral directly from equations (1). We multiply the first two of equations (1) by 1 and $\pm i$, respectively; there then follows by addition

$$(5') \qquad A\frac{d(p \pm iq)}{dt} = \pm i(C-A)r(p \pm iq),$$

or

$$(6') \qquad p \pm iq = (p_0 \pm iq_0)e^{\pm\frac{(C-A)r_0}{A}it},$$

where p_0, q_0, and r_0 denote the values of p, q, and r at time $t = 0$.

*) Cf. Sur la rotation d'un corps, Crelle Journal Bd. 39, 1850.[73]

*) Sur quelques applications des fonctions elliptiques, Paris 1883.[74]

It remains to integrate equations (2). From the integrals of these equations contained in (7), we now conclude

$$(7') \qquad \cos\vartheta = \frac{Cr}{G} = \text{const.},$$

$$\sin\vartheta\, e^{\pm i\varphi} = \pm i\frac{A}{G}(p \mp iq).$$

From the last equation, there follows through logarithmic differentiation, with consideration of (5′) and (7′),

$$\frac{d\varphi}{dt} = -\frac{C-A}{A}r = \text{const.},$$

or

$$(8') \qquad \varphi = \mu t,$$

where we set, as an abbreviation,

$$(8'') \qquad \mu = \frac{A-C}{A}r$$

and assume that the point of time $t = 0$ corresponds to the angle $\varphi = 0$.

Finally, we have, according to (8),

$$\frac{d\psi}{dt} = \frac{G}{A} = \text{const.},$$

and therefore

$$(9') \qquad \psi = \nu t,$$

where we again set, as an abbreviation,

$$(9'') \qquad \nu = \frac{G}{A}$$

and assume that $\psi = 0$ for $t = 0$. *Through equations* (7′), (8′), *and* (9′), *the general motion of the force-free symmetric top is characterized as regular precession.*

We will regard, as previously, the quantities μ, ν, and ϑ as the primary constants of the regular precession. According to equations (4) of page 50, the constants p_0, q_0, r_0 may be expressed in terms of μ, ν, and ϑ as

$$(10) \qquad p_0 = 0, \quad q_0 = \nu\sin\vartheta, \quad r_0 = \mu + \nu\cos\vartheta.$$

On the other hand, however, the three quantities μ, ν, and ϑ are equivalent to the two earlier integration constants h and G; there is thus a relation between μ, ν, and ϑ, which we wish to establish. According to equation (8″), we have

$$-\mu A = (C-A)r = (C-A)r_0,$$

and thus, with consideration of (10),

$$-\mu A = (C - A)(\mu + \nu \cos\vartheta),$$

or

$$(11) \qquad C\mu + (C - A)\nu \cos\vartheta = 0.$$

For a particular force-free symmetric top, only one specific type of regular precession can occur, in which the constants μ, ν, ϑ satisfy the relation (11).

Finally, we wish to classify the regular precession of the symmetric top according to the value of $\frac{\nu}{\mu}$, in the sense of page 53. Equation (11) yields

$$\frac{\nu}{\mu} = \frac{C}{A - C}\frac{1}{\cos\vartheta}.$$

Since we may assume, in general, that the angle ϑ is smaller than $\frac{\pi}{2}$ (cf. page 51), the right-hand side is positive if $A > C$ and negative if $A < C$. As a result, we can say, on the basis of the table of page 53,

The regular precession in question is progressive for the prolate top $(A > C)$ *and retrograde for the oblate top* $(A < C)$.

In the latter case, the value of $\frac{\nu}{\mu}$ implies a still more precise limitation. Namely, we have, if $C > A$,

$$-\cos\vartheta\,\frac{\nu}{\mu} = \frac{C}{C - A} > 1,$$

or

$$\frac{\nu}{\mu} < -\frac{1}{\cos\vartheta}.$$

For the oblate top $(A < C)$*, the "free" precession is therefore necessarily pericycloidal, and for the prolate top* $(A > C)$ *epicycloidal.*

Of the four intervals for the value of $\frac{\nu}{\mu}$ distinguished on page 53, only the outer two, therefore, come into consideration for the force-free motion of the top. In contrast, we will see in the consideration of the heavy top (cf. the sixth section of this chapter) that the two middle intervals are also accessible for its precessional motion.

In the limiting case $A = C$,

$$\frac{\nu}{\mu} = \pm\infty;$$

the top rotates, according to page 54, uniformly about the vertical; that

is, about the impulse axis. We arrive therefore at the well-known result:

For the spherical top $A = C$, uniform precession is transformed into a uniform rotation about the axis of the impulse.

§3. On the meaning of the Euler equations and their relation to the equations of Lagrange

The Euler equations occupy an entirely singular position in the system of mechanics, and are not classified under the general type of differential equation for mechanics that was established by L a g r a n g e. In addition, it is not possible to establish equations for arbitrary mechanical systems that offer advantages similar to those of the Euler equations for the rigid body. It will thus be interesting to discuss, in this section, the basis of the uniqueness of the Euler equations, and to clarify their advantages compared with the general equations of Lagrange.

We first present the general Lagrange equations for the case of the top. We imagine the position of the top to be represented by the three position coordinates φ, ψ, ϑ, and the state of velocity to be represented by the corresponding velocity coordinates φ', ψ', ϑ'. We understand by T the expression for the *vis viva* written in the given position and velocity coordinates, and by dA the work that the system of external forces applied to the top performs in the infinitely small displacement $(d\varphi, d\psi, d\vartheta)$.

The equations now take precisely the same form that we recognize from the preceding chapter for the single mass particle (cf. page 79 and ff.); namely, we have, on the one hand,

$$(1) \qquad \begin{cases} \dfrac{d[\Phi]}{dt} - \dfrac{\partial T}{\partial \varphi} = \Phi, \\ \dfrac{d[\Psi]}{dt} - \dfrac{\partial T}{\partial \psi} = \Psi, \\ \dfrac{d[\Theta]}{dt} - \dfrac{\partial T}{\partial \vartheta} = \Theta, \end{cases}$$

and, on the other hand,

$$(2) \qquad \begin{cases} [\Phi] = \dfrac{\partial T}{\partial \varphi'}, \quad [\Psi] = \dfrac{\partial T}{\partial \psi'}, \quad [\Theta] = \dfrac{\partial T}{\partial \vartheta'}, \\ \Phi = \dfrac{\partial A}{\partial \varphi}, \quad \Psi = \dfrac{\partial A}{\partial \psi}, \quad \Theta = \dfrac{\partial A}{\partial \vartheta}. \end{cases}$$

In fact, one always has, according to Lagrange, the same system of equations for any system of three degrees of freedom, however one may introduce the three parameters φ, ψ, ϑ, and also the corresponding system of equations for a system of n degrees of freedom, where n parameters are used instead of our three parameters.

Our manner of writing the Lagrange equations differs from the usual manner of writing only in so far as we express the concept of the impulse explicitly. This is recommended throughout, since the meaning of the equations becomes much clearer in this way. Equations (2) evidently determine the components of the impulse that pertain to a given position of the top and a given state of velocity, as well as the components of the external force that correspond to a given position. (The differential symbols $\dfrac{\partial A}{\partial \varphi}\cdots$ are understood in the same uncharacteristic sense as explained above on page 77.) Equations (1) then determine how the components of the impulse change during the time element dt. Equations (1) are thus, in the case of the top, equivalent to the Euler equations, and must express the fact that the change of the impulse is equal, at each time, to the turning-impact produced by the external forces.

We wish here to actually carry out the conversion of the Euler equations into the Lagrangian form, and thus give a derivation of the latter from the impulse principle for the present case. Since the calculations become somewhat detailed, however, we will restrict ourselves here to the case of the symmetric top.

We first express the rotation vector (p, q, r) in terms of the velocity coordinates φ', ψ', ϑ', and express, in a corresponding manner, the components $[\mathsf{\Phi}]$, $[\mathsf{\Psi}]$, $[\mathsf{\Theta}]$ and $\mathsf{\Phi}$, $\mathsf{\Psi}$, $\mathsf{\Theta}$ of the impulse and the external force appearing in the Lagrange equations in terms of the components L, M, N and $\mathsf{\Lambda}$, M, N appearing in the Euler equations. This is done, according to equations (7) of page 45 and page 108, in the following manner:

$$(3)\qquad \begin{cases} p = \sin\varphi\,\sin\vartheta\cdot\psi' + \cos\varphi\cdot\vartheta', \\ q = \cos\varphi\,\sin\vartheta\cdot\psi' - \sin\varphi\cdot\vartheta', \\ r = \varphi' + \cos\vartheta\cdot\psi'; \end{cases}$$

$$(4)\qquad \begin{cases} [\mathsf{\Phi}] = N, \\ [\mathsf{\Psi}] = \sin\vartheta\,\sin\varphi\cdot L + \sin\vartheta\,\cos\varphi\cdot M + \cos\vartheta\cdot N, \\ [\mathsf{\Theta}] = \cos\varphi\cdot L - \sin\varphi\cdot M; \end{cases}$$

$$(5)\qquad \begin{cases} \mathsf{\Phi} = & & & \mathsf{N}, \\ \mathsf{\Psi} = \sin\vartheta\sin\varphi\cdot\mathsf{\Lambda} + \sin\vartheta\cos\varphi\cdot\mathsf{M} + \cos\vartheta\cdot\mathsf{N}, \\ \mathsf{\Theta} = \cos\varphi\cdot\mathsf{\Lambda} - \sin\varphi\cdot\mathsf{M}. \end{cases}$$

We will, further, have need of the expressions for the *vis viva* and the work, which, according to pages 108 and 109, respectively, take the form

$$(6)\quad T = \frac{1}{2}(A(p^2+q^2)+Cr^2) = \frac{A}{2}(\sin^2\vartheta\cdot\psi'^2+\vartheta'^2)+\frac{C}{2}(\varphi'+\cos\vartheta\cdot\psi')^2,$$

$$(6')\qquad dA = (\mathsf{\Lambda}p + \mathsf{M}q + \mathsf{N}r)\,dt = \mathsf{\Phi}\,d\varphi + \mathsf{\Psi}\,d\psi + \mathsf{\Theta}\,d\vartheta.$$

From equations (4) through (6′), we immediately infer the existence of equations (2); these equations are therefore an immediate consequence of the previous definitions of the impulse components $[\mathsf{\Phi}]$, $[\mathsf{\Psi}]$, $[\mathsf{\Theta}]$ and the force components $\mathsf{\Phi}$, $\mathsf{\Psi}$, $\mathsf{\Theta}$.

We now take the Euler equations in their form (3′) on page 141, after we have set $A = B$. We first see that the third Euler equation

$$\frac{dN}{dt} = \mathsf{N}$$

is directly identical with the Lagrange equation

$$\frac{d[\mathsf{\Phi}]}{dt} - \frac{\partial T}{\partial\varphi} = \mathsf{\Phi}$$

corresponding to the φ coordinate. In fact, we have, according to (4), (5), and (6),

$$N = [\mathsf{\Phi}],\qquad \mathsf{N} = \mathsf{\Phi},\qquad \frac{\partial T}{\partial\varphi} = 0.$$

To obtain the Lagrange equation corresponding to the ψ coordinate, we multiply the Euler equations sequentially by the coefficients of L, M, N in the expression (4) for $[\mathsf{\Psi}]$. There follows, according to (4) and (5),

$$(7)\qquad \begin{cases} \sin\vartheta\sin\varphi\dfrac{dL}{dt} + \sin\vartheta\cos\varphi\dfrac{dM}{dt} + \cos\vartheta\dfrac{dN}{dt} \\ = \dfrac{d[\mathsf{\Psi}]}{dt} - \dfrac{d\sin\vartheta\sin\varphi}{dt}L - \dfrac{d\sin\vartheta\cos\varphi}{dt}M - \dfrac{d\cos\vartheta}{dt}N \\ = (A-C)\sin\vartheta(\sin\varphi\cdot q - \cos\varphi\cdot p)r + \mathsf{\Psi}. \end{cases}$$

But one now recognizes, if one sets $L = Ap$, $M = Aq$, $N = Cr$ and applies equations (3), that

$$\frac{d\sin\vartheta\sin\varphi}{dt}L + \frac{d\sin\vartheta\cos\varphi}{dt}M + \frac{d\cos\vartheta}{dt}N$$
$$= (C-A)\sin\vartheta(\sin\varphi\cdot q - \cos\varphi\cdot p)r.$$

Thus there follows simply from (7)

$$\frac{d[\Psi]}{dt} = \Psi.$$

But this equation is identical with the second Lagrange equation for the top problem, since, according to (6),

$$\frac{\partial T}{\partial \psi} = 0.$$

To derive, finally, the equation corresponding to the ϑ coordinate from our Euler equations, we proceed in the same way. We multiply the Euler equations sequentially by $\cos\varphi$, $-\sin\varphi$, and 0; that is, by the coefficients of L, M, and N in the third of equations (4). We thus obtain

$$\text{(8)}\qquad \begin{cases} \cos\varphi \dfrac{dL}{dt} - \sin\varphi \dfrac{dM}{dt} = \dfrac{d[\Theta]}{dt} - \dfrac{d\cos\varphi}{dt} L + \dfrac{d\sin\varphi}{dt} M \\ \qquad\qquad = (A - C)(\cos\varphi \cdot q + \sin\varphi \cdot p) r + \Theta. \end{cases}$$

But now according to (3),

$$-\frac{d\cos\varphi}{dt} L + \frac{d\sin\varphi}{dt} M = A(\sin\vartheta \cdot p + \cos\vartheta \cdot q)\varphi' = A \sin\vartheta \cdot \psi' \, \varphi',$$

and, on the other hand,

$$(A - C)(\cos\varphi \cdot q + \sin\varphi \cdot p) r = (A - C) \sin\vartheta \cdot \psi'(\varphi' + \cos\vartheta \cdot \psi').$$

As a result, we can write, instead of (8),

$$\frac{d[\Theta]}{dt} - \Big(A \cos\vartheta \sin\vartheta \cdot \psi'^2 - C \sin\vartheta \cdot \psi'(\varphi' + \cos\vartheta \cdot \psi')\Big) = \Theta.$$

But this, with consideration of the expression for the *vis viva*, is again nothing other than the third Lagrange equation.

In addition to the already discussed Lagrange equations (or the equations "of the second kind"), one also uses in analytic mechanics, as is well known, the so-called *Lagrange equations of the first kind.* For a system of n particles and k degrees of freedom, these consist of $3n$ differential equations for the coordinates of the points, in which appear $3n - k$ disposable parameters, called *Lagrange multipliers*, which are determined so that the coordinate changes defined by the differential equations are compatible with the characteristic freedom of motion of the system.

Now we cannot, to be sure, think of using the Lagrange equations of the first kind directly in the case of the top. We would have—corresponding to the infinitely many mass particles of which the top

consists—to write infinitely many equations with infinitely many Lagrange multipliers. But we may indeed produce differential equations for the top that stand, so to speak, in the middle ground between the Lagrange equations of the first and second kinds, and which we may denote as *Lagrange equations of mixed type.*

The common feature that binds these mixed equations with the equations of the first kind is the introduction of *supernumerary coordinates.* One determines the position of the system through more coordinates than the number of degrees of freedom: if one uses the $3n$ rectangular coordinates of the n particles of the system as supernumerary coordinates, then one arrives, as we said, at the Lagrange equations of the first kind. If one uses, in contrast, $3n-1$, $3n-2, \dots, k+1$ coordinates, then there results each time equations of our mixed type. If one uses only the minimum number of k coordinates, the Lagrange equations of the second type result.

Corresponding to the supernumerary position coordinates, one introduces as supernumerary velocity coordinates the differential quotients of the position coordinates with respect to time. But it is also further necessary to define the force and impulse coordinates corresponding to the velocity coordinates. For this one will have to begin, as earlier (cf. page 92), from the expression for the work of the system and for its *vis viva*, respectively. But these expressions are not completely determined in their dependence on the supernumerary coordinates, but rather can be modified formally through the addition of arbitrary multiples of the relations existing between them. As a result, the force and impulse coordinates in question remain, in a certain sense, undetermined; they contain arbitrary parameters, whose appearance, from our standpoint, explains the necessity of the Lagrange multipliers. The equations that determine the changes of these supernumerary impulse coordinates may be written, moreover, according to the schema of the general Lagrange equations.

The differential equations of the top written in terms of the parameters α, β, γ, δ or the quaternion parameters A, B, C, D obviously belong to this mixed type. In fact, we use supernumerary coordinates when we take these parameters as a basis for the determination of the position and velocity states. The resulting differential equations will contain a Lagrange multiplier which is determined in such a way that

the changes of α, β, γ, δ or A, B, C, D that follow from the differential equations are always compatible with the condition

$$\alpha\delta - \beta\gamma = 1 \quad \text{or} \quad A^2 + B^2 + C^2 + D^2 = 1.$$

Moreover, these differential equation take a truly clear and elegant form. We will nevertheless forgo their development in this place, since we would otherwise have to go more deeply into the definition of the supernumerary impulse coordinates.[75] —

We now return to the exercise posed at the beginning of this section: to study the uniqueness of the Euler equations compared with the equations of Lagrange.

It is easy to see why the Euler equations differ formally from the Lagrange equations. In the Euler equations, we determine the velocity state in terms of the components of the rotation vector p, q, r with respect to the axes X, Y, Z. These coordinates are, as emphasized on page 46, "inexact velocity coordinates"; that is, they are not differential quotients with respect to time of any spatial measures. Now the Lagrangian form of the mechanical differential equations indeed persists, as stressed on page 155, for the introduction of arbitrary position coordinates. But the velocity coordinates are then necessarily the derivatives of the chosen position coordinates with respect to time, and are therefore always exact differential quotients. Thus we can say: *the Euler equations are not a special case of the Lagrange equations, in so far as we determine the velocity state in the Euler equations with inexact coordinates, while the Lagrange equations assume the use of exact velocity coordinates.*

But we will ask, further, for the basis of the special merits of the Euler equations compared with the general differential equations of mechanics. These appear primarily for the force-free motion of the top, to which we will first refer.

We know from the previous chapter, and it is also self-evident from the start, that the determination of the impulse relative to the top depends at each instant only on the instantaneous velocity state of the top, and is completely independent of the position of the top in space. In fact, we could determine the position of the impulse in the body analytically in terms of the differential quotients of the expression for the *vis viva* that depends only on the components p, q, r of the velocity

state, or also geometrically by the construction that uses the ellipsoid of inertia, in which the position of the ellipsoid of inertia in space comes in no way into the question. We therefore always obtain the same impulse vector relative to the top (for the same position and magnitude of the rotation vector), whatever position we may give to the top in space.

Further, the rate of change of the impulse relative to the top is determined, for force-free motion, only by the positions of the impulse and rotation vectors relative to the top and the magnitudes of these vectors, and depends in no way on the absolute position of the top in space, as follows immediately from the theorem of the permanence of the impulse.

The determination of the impulse and the impulse change, or, what signifies the same thing, the rotation vector and the change of the rotation vector relative to the top, is therefore a problem, for force-free motion, in which the position of the top does not come into question. The problem remains the same, however we may choose the initial position of the top in space.

One may suppose from the beginning that an analytic formulation of our problem is possible which exhibits its independence of the absolute position of the top in space, and in which, therefore, the position coordinates of the top do not appear at all. *This formulation is now realized directly by the Euler equations for force-free motion.* In fact, only the quantities p, q, r (or L, M, N) and their differential quotients with respect to time appear in these equations, which have a meaning independent of the position of the top.

In the Lagrange equations, in contrast, we decompose the rotation vector by use of the position coordinates φ, ψ, ϑ into the components φ', ψ', ϑ', which refer to components on axes that are not fixed in the top, so that the corresponding values of φ', ψ', ϑ' will be changed for a change in the absolute position of the top. As a result, the Lagrange equations do not, without further information, give prominence to the fact that determination of the rotation vector and its position with respect to the top is independent of the position of the top in space.

However, the independence of the Euler equations from the position coordinates brings with it (for force-free motion) a considerable *simplification of the integration procedure.* Namely, we can first, as is done in the previous section, integrate the differential equations for p, q, r

in themselves, and thus settle all questions that refer to the positions of the impulse and rotation vectors with respect to the top. The determination of the absolute position in space then demands, as a second step, the integration of the "kinematic" equations, where we still have full freedom in the choice of the position coordinates.

A corresponding decomposition of the integration difficulties is not possible for the Lagrange equations, unless one directly lays as a basis for the integration those linear functions of φ', ψ', ϑ' that are equal, respectively, to p, q, r. But this would mean nothing other than the transition after the fact to the Euler equations.

If external forces influence the motion of the top, it is generally illusory to speak of the advantage that the Euler equations offer over the Lagrange equations. Namely, since the external forces will depend on the position of the top in space, the change of the impulse vector will also no longer be independent of the position. *As a result, the Euler equations form* (unless special symmetry relations concerning the external forces exist, into which we cannot go in this place), *with the kinematic equations, an irreducible simultaneous system of differential equations.* We can, therefore, as will be done in the following chapter, begin the integration just as well from the Lagrange equations as from the Euler equations.

The preceding considerations can be made shorter and more precise if we allow ourselves to assume a few concepts from the *Lie theory of transformation groups.*[76] Namely, we can then say that the problem of determining the magnitude and position of the rotation vector for force-free motion of the top admits of a triply infinite group of transformation; namely, the collected rotations of the coordinate system x, y, z about the point O*). Here the components p, q, r play the role of invariants of the group. The decomposition of the mechanical differential equations into one system of equations (the Euler equations) that behave invariantly with respect to transformations of the group, and a second system of equations (our kinematic equations) that are changed by these transformations and directly represent the dependence of the top on the system x, y, z, is anticipated by the Lie theory.

---

*) More complete details regarding this are found in Mr. H. Liebmann: Bemerkungen über integrable Fälle der Kreiselbewegung, Math. Ann. Bd. 50.[77]

With the help of this theory, it is possible to answer the general question of when, for other mechanical systems, equations are possible that are analogous to the Euler equations for the rigid body, and that bring the advantage of the division of the integration process into two simpler steps. Such equations will be given only in the special cases in which the mechanical system and the applied external forces admit of an appropriate group of continuous transformations. *In general, it will thus not be possible to speak of an analogue to the Euler equations.*

If we reverse our manner of consideration and view the existence or nonexistence of the Euler equations as equivalent to the existence or nonexistence of an impulse theorem, then we can complement our last claim by the following statement, into whose basis we will enter more deeply in later chapters: *For general mechanical systems, it will not be possible to speak of a direct analogue to the impulse theorems for the rigid body.*

§4. Guidance of the top on a prescribed path. D'Alembert's principle

After having dispatched the "natural" motion of the force-free top in the second section of this chapter, we could now inquire of the motion that the top executes on the basis of given forces, and, in particular, the force of gravity. Instead, we will first pose a simpler exercise, and consider the phenomena associated with an arbitrary "enforced" motion.

We wish to suppose that we somehow move the top in a prescribed way with our hand, in such a way that a straight line fixed in the top describes a cone with its vertex at O, and, moreover, let the top rotate in an arbitrarily given manner about this straight line. We will first assume that external forces, other than those exerted by our hand, are not present. We will therefore, in particular, let the center of gravity coincide with the support point.

A certain force will be necessary at each moment for the generation of our enforced motion. We have the feeling that the top applies a resistance against the guidance, and sense, correspondingly, a pressure on our hand. We can, in general, call this resistance the *inertial resistance.*

We consider, in the following, the determination of the inertial resistance in magnitude and direction.

To form a connection with something known, we wish to begin with the corresponding question for a single mass particle. We therefore imagine an otherwise force-free particle of mass m guided along an arbitrary curve with an arbitrary velocity, and ask for the resistance with which the particle opposes this motion.

We first construct the impulse of the particle, which we denote by i, at each position of the path. If the motion is not directly the "natural motion," that is, uniform progression in a straight line, then the impulse will change from position to position. But now, according to Newton's lex secunda, each change di of the impulse requires a force P of such a nature that $P\,dt = di$. We must exert this force with our hand; the opposed equal force is the resistance W that we have to overcome, and which we call the *inertial resistance.* Thus

$$W = -\frac{di}{dt}. \tag{1}$$

We can thus determine the resistance in question geometrically as follows. We construct, for two neighboring time points t_0 and t_1, the impulses of the forced motion, bring the initial points of the two vectors into coincidence by a parallel displacement, and complete them in a triangle. Then the third side of the triangle, divided by $t_0 - t_1$, yields the desired resistance.

If we wish to compute the components W_x, W_y, W_z of our resistance with respect to three rectangular coordinates axes, these follow, naturally, if we decompose equation (1) into components. But we can read off these components of W directly from our earlier equations of motion for the single mass particle. In the construction of those equations (see page 77), we thought of the external force (X, Y, Z) as given, and asked for the "natural motion" that ensues under its influence. But we can also think of the motion, that is, the successive changes of the impulse, as somehow given, and ask for the corresponding force that we must exert to enforce the considered motion. The resistance in question is equal to the opposite of this force. We thus conclude, from the cited equations (9) on page 77, that

$$(2) \qquad W_x = -X = -\frac{d[X]}{dt}, \quad W_y = -Y = -\frac{d[Y]}{dt}, \quad W_z = -Z = -\frac{d[Z]}{dt}.$$

This evidently coincides with equation (1).

If an external force P acts on our mass particle, then the force required for the generation of the enforced motion will obviously be reduced by its amount. Correspondingly, the resistance that we then have to overcome, and which we denote by W', will be given, instead of by (1), by the equation

$$(3) \qquad W' = -\frac{di}{dt} + P.$$

Its components result from the equations of motion of the single mass particle, if we denote the total force required for the generation of the motion by $(X + X',\ Y + Y',\ Z + Z')$, where (X, Y, Z) denote the components of the externally acting force P and (X', Y', Z') denote the components of the force applied for the guidance of the particle, as follows:

$$(4) \qquad \begin{cases} W'_x = -\,X' = -\,\dfrac{d[X]}{dt} + X, \\ W'_y = -\,Y' = -\,\dfrac{d[Y]}{dt} + Y, \\ W'_z = -\,Z' = -\,\dfrac{d[Z]}{dt} + Z. \end{cases}$$

We now carry over the preceding remarks, entirely self-evident according to the fundamental laws of general mechanics, to the case of the *generalized force-free top.*

We first construct the impulse for each stage of the motion, and thus obtain a generally variable vector i emanating from O. But each change di of the impulse requires, as we know from the transference of the second Newtonian law of motion given on page 115, a turning-force D of such a nature that $D\,dt = di$. We must exert the turning-force with our hand for the guidance of the top; the top resists the motion with an equal oppositely directed force. *The desired resistance is therefore given in direction, magnitude, and sense by the vector equation*

$$(1') \qquad W = -\frac{di}{dt}.$$

We find the resistance once more geometrically by constructing the impulse in space for two neighboring time points t_0 and t_1, and binding

the endpoints of the two vectors. The ratio of the binding line to the time interval $t_0 - t_1$ then yields the magnitude, direction, and sense of W at the time point $t_0 = t_1$.

To compute the components W_X, W_Y, W_Z of our resistance with respect to the coordinate system XYZ fixed in the top, we draw upon the Euler equations.

In fact, we can conceive of these equations, as above for the equations of motion of a single mass particle, and in the same way for the differential equations of mechanics in general, in two ways. We can, on the one hand, think of the external forces as given and ask for the motion that the system would execute from a given initial state on the basis of these forces. But we can also, on the other hand, think of the motion of the system as arbitrarily given, and ask for the forces that are required for the generation of this motion. While the former exercise necessitates the *integration* of the mechanical equations of motion, the latter is settled through the *differentiation* signified in the equations. In this section we consider the Euler equations characterized from the latter standpoint.

The desired inertial resistance is obviously, according to the law of equality of action and reaction, equal and oppositely directed to the force necessary for the generation of the motion. If we therefore take the components $\mathsf{\Lambda}$, M, N of this force from the Euler equations of page 140, then we obtain for the components of the resistance, referred to the coordinate system fixed in the body, the following values:

$$
(2') \qquad \begin{cases} W_X = -\,\mathsf{\Lambda} = -\dfrac{dL}{dt} + Mr - Nq, \\ W_Y = -\,\mathsf{M} = -\dfrac{dM}{dt} + Np - Lr, \\ W_Z = -\,\mathsf{N} = -\dfrac{dN}{dt} + Lq - Mp. \end{cases}
$$

It is evident from our derivation of the Euler equations that this determination of W coincides with the value contained in $(1')$.

We now wish to assume, further, that a somehow given force system acts on our top, which, composed in the usual way, results in a turning-force D. Let its components with respect to the coordinate axes be $\mathsf{\Lambda}$, M, N.

Then a part of the impulse change will be compensated by this turning-force D. The remaining uncompensated part gives, taken negatively

in direction, the magnitude and sense of the resistance that we must overcome for the enforced motion of the top. This resistance W' now differs from the inertial resistance W of the top by the amount of the turning-force D, and is correspondingly calculated from the vector equation

$$W' - W = D \quad \text{or} \quad W' = -\frac{di}{dt} + D. \tag{3'}$$

Its components with respect to the coordinate axes follow once more from the Euler equations. If we denote the components of the total force that is required for the generation of the motion in question by $\mathsf{\Lambda} + \mathsf{\Lambda}'$, $\mathsf{M} + \mathsf{M}'$, $\mathsf{N} + \mathsf{N}'$, then we obtain

$$\left\{\begin{aligned} W'_X &= -\mathsf{\Lambda}' = -\frac{dL}{dt} + Mr - Nq + \mathsf{\Lambda}, \\ W'_Y &= -\mathsf{M}' = -\frac{dM}{dt} + Np - Lr + \mathsf{M}, \\ W'_Z &= -\mathsf{N}' = -\frac{dN}{dt} + Lq - Mp + \mathsf{N}. \end{aligned}\right. \tag{4'}$$

From these equations we attain a very simple and very well known *criterion by which we can characterize the* n a t u r a l *motion of the top occurring under the influence of given forces, in contrast to an arbitrary* e n f o r c e d *motion.* We wish here also to speak first of a single mass particle. The natural motion which the particle executes under the influence of the force P without external constraint is obviously that motion for which the previously introduced resistance W' vanishes. *The condition for the natural motion is therefore, according to* (3),

$$-\frac{di}{dt} + P = 0. \tag{5}$$

This equation is, fundamentally, completely trivial; it is nothing other than the basic impulse theorem on which, as the second axiom, the mechanics of Newton is based.

But we can now give this theorem a new formulation in which we reduce the question of the natural motion of a particle to a question of equilibrium. This formulation is suggested as soon as one, as done above, introduces the negatively taken rate of change of the impulse as a special force, and gives it a special name ("inertial resistance"). Namely, we can conceive of equation (5) in words in the following manner:

In the natural motion of a particle, the inertial resistance of the particle always maintains equilibrium with the external force.

Naturally, nothing is achieved with this formulation in the preceding simplest case. We mention this only because the corresponding statement governs the natural motion of arbitrary mechanical systems as the so-called *D'Alembert's principle*, to which we will soon return.

We first characterize the natural motion of the generalized top in the corresponding way. This is, obviously, again the motion against which the top exerts no resistance, and for which, therefore, W' in equation $(3')$ vanishes. *Our criterion for the natural motion of the top occurring under the influence of the external turning-force D is thus completely similar to that for the particle:*

$$-\frac{di}{dt} + D = 0. \tag{5'}$$

As a basis, we fall back again on our fundamental impulse theorem of page 115, or on the equivalent Euler equations.

To formulate this equation in the present consideration, we return to the individual mass particles that constitute the top, and imagine the total impulse i and the resulting turning-force D into which the individual impulses and individual forces belonging to the mass particles are resolved. We can then state equation $(5')$ as an equilibrium condition in the sense of elementary statics, and can formulate the following theorem, which coincides with D'Alembert's principle in the case of the top:

In the natural motion, the system of the inertial resistances of all the individual mass elements is, due to the rigid binding of the mass particles, in equilibrium with the system of externally applied forces on the particles.[78]

This statement now carries over, as stated, to arbitrary mechanical systems with arbitrary (not necessarily rigid) bindings. All questions of motion thus appear reduced to pure equilibrium questions. In this static conception of the dynamic laws lies the true accomplishment of D'Alembert and the character of his principle.

Finally, we can consider, so to speak, an *intermediary between the natural and the enforced motion of the top.* We wish, to have a definite

case in view, to let a principal axis H of the top describe a cone with vertex at O in a specified way, or, in particular, be thought fixed, while we do not influence the rotation about this axis by external compulsion. The latter part of the motion is then designated as a natural motion, and the former as an enforced motion. The components of the resistance W' in the direction of H must vanish, since these cannot be compensated through the guidance of H. In contrast, the components of W' perpendicular to H will have a nonzero value, which signifies to us exactly the applied pressure for the guidance. Correspondingly, only the components of W' perpendicular to H are to be computed according to equation $(3')$.

The manner of expression of D'Alembert's principle also holds perfectly for this type of motion. Exactly in this versatility, one may say, rests the essential usefulness of the principle.

If we regard that part of the motion which we just designated as the natural motion, then we have, so to speak, a *top with only one natural motion degree of freedom.* In this case, the inertial resistances of the individual mass particle evidently maintain equilibrium with the components of equation $(5')$ which come into consideration for us; that is, those resistances and forces which, taken together, produce no rotation about the axis H. But if we also bring the enforced motion of H into consideration, then we have a *top with one natural motion degree of freedom and two enforced motion degrees of freedom.* The equilibrium between the inertial resistances and the exterior forces will be established here if we add to the external forces the force that we must exert for the guidance of the top. The resulting manner of consideration is thus preferred to the previous, because it admits of the determination of not only the course of the natural motion, but also the pressure necessary for the enforced motion.

In the usual language of analytic mechanics, one would formulate the preceding restriction of the natural motion by imposing on the top *certain constraint equations*, which in the above example (for arbitrary guidance of the axis H) would still contain the time. The usefulness of D'Alembert's principle consists, analytically speaking, exactly in the fact that one can directly eliminate, by means of this principle, as many such constraint equations as one wishes, replacing them in the sequel by appropriate forces.

In the last section of this chapter, we will take up some examples of the application, here only indicated, of D'Alembert's principle to motions of a partially enforced character.

§5. Special development for the spherical top. Decomposition of the total resistance into an acceleration resistance and a deviation resistance

In particle mechanics, the consideration of the inertial resistance with which a single mass particle opposes an enforced motion can be taken somewhat farther than was done in the previous section.

We wish, as usual, to project the resistance onto the tangent to the particle trajectory, and thus decompose the inertial resistance into two components. One component, which may be called the *tangential force*, falls in the direction of the tangent; the second component, which is called the *centrifugal force* and which already appeared in the first section of this chapter, is perpendicular to the tangent. We wish to denote the magnitude of the tangential force by H, and the magnitude of the centrifugal force by K.

We can also define the centrifugal force as *the component of the resistance whose overcoming requires no work.* Since, according to its definition, the centrifugal force is perpendicular to the instantaneous direction of the motion, there follows, if we denote its components with respect to some coordinate axes by K_x, K_y, K_z,

$$x'K_x + y'K_y + z'K_z = 0;$$

on the other hand, the overcoming of the centrifugal force requires in the time dt the work

$$dA = (x'K_x + y'K_y + z'K_z)\,dt.$$

This work therefore vanishes, as claimed.

To arrange the calculation of H and K as simply as possible, one proceeds, as is well known, by replacing the given motion by another, which approximates the given motion at the considered point of time to the second order; that is, which, for that point of time and a point immediately following, has the same magnitude and direction of the impulse, and thus also the same tangential and centrifugal forces as the original motion. One constructs, namely, the curvature circle to the given trajectory at the considered place, and lets the mass particle m traverse this circle with uniform acceleration, in such a way that its

velocity and acceleration coincide with the velocity and acceleration of the original motion at the position in question.

One then easily finds, from equation (1) of the preceding section, the *magnitude of the tangential force*:

$$H = m\frac{dv}{dt} = \frac{d|i|}{dt}\,{}^{*}). \tag{1}$$

The direction of the force is, according to the original definition, identical with the direction of the instantaneous velocity. *The sense* is opposite to the sense of the instantaneous acceleration.

In the same manner, there follows for the *magnitude of the centrifugal force* the previously used formula

$$K = \frac{mv^2}{R}, \tag{2}$$

in which R is the radius of the curvature circle. The *direction of the centrifugal force* is determined so that it 1) stands perpendicular to the instantaneous direction of motion and 2) lies in the osculating plane of the trajectory, and therefore in the plane of the curvature circle. We can establish the *sense of the centrifugal force* through the following rule: it seeks to distance the mass particle from the center of the curvature circle.

According to (1), the tangential force vanishes only when the acceleration vanishes. On the other hand, the centrifugal force vanishes, according to (2), if the curvature of the path is equal to zero, and therefore when the velocity has an invariable direction. As a result, we can claim that *the appearance of the tangential force has its basis in a change of length of the velocity vector, and the centrifugal force has its basis in a change of direction of the velocity vector.*

If we now wish, in the case of the generalized top, to effect a similar decomposition of the inertial resistance into two perpendicular components, difficulties appear at once. We must rather descend, in order that the concepts of the tangential and centrifugal forces can carry over smoothly, to the special case of the *spherical top.* It is then verified, as already remarked on page 134, that the spherical top (by virtue of the coincidence of the impulse and velocity axes) presents, in kinetic respects, a complete analogy with the single mass particle. This

---

*) Concerning the meaning of the symbol $|i|$ as the expression for the length of the vector i, cf. the footnote on page 143.

restriction of generality, moreover, signifies little for the intended further developments, since we will, in any case, turn our primary interest in the following chapters toward the spherical top.

The inertial resistance W of the spherical top is naturally given, just as for the generalized top, by equation $(1')$ of the preceding section. We now project W perpendicularly onto the axis of the instantaneous rotation, and thus obtain a turning-force that has the instantaneous rotation vector as its axis, and which we will call the *acceleration resistance*, and a second turning-force, whose axis stands perpendicular to the rotation vector, and which shall be called the *deviation resistance.* The basis for these designations will be clear directly. We will denote the magnitude of the acceleration resistance by H, and that of the deviation resistance by K.

We first mention an important characteristic of the deviation resistance, which this resistance has in common with the centrifugal force: *no performance of work is necessary for the overcoming of the deviation resistance.* In fact, the definition of the deviation resistance states that its axis is continuously perpendicular to the instantaneous rotation axis p, q, r. If we denote the components of our resistance by K_X, K_Y, K_Z with respect to the same axes to which p, q, r are referred, we therefore have

$$pK_X + qK_Y + rK_Z = 0.$$

As a result, the necessary work for the overcoming of the deviation resistance during the time interval dt becomes

$$dA = (pK_X + qK_Y + rK_Z)\,dt = 0.$$

For the calculation of the acceleration and the deviation resistance, we will now proceed as above for the single mass particle. Namely, we construct, in addition to the given motion, a simpler motion, which approximates the given motion to the second order at the position in question. According to page 54, this is achieved by a properly chosen *"uniformly accelerated precession."* We will effect this motion if we lay, on the given polhode and herpolhode cones, the curvature cones that osculate on the instantaneous rotation axis, and mark on the curvature cones of our precessional motion the spiral polhode and herpolhode curves that are tangent, respectively, to the polhode and herpolhode curves of the original motion at the endpoint of the instantaneous rotation vector. This new motion has, at two neighboring moments of time, the same rotation vector in direction and magnitude, and therefore also

the same impulse vector, as the given motion. As a result, the acceleration and deviation resistances of the two motions will also be the same.

For the calculation of H and K, we must now consider the displacement di that the endpoint of the impulse vector experiences in space during the time dt for our uniformly accelerated precession. If we denote by ds the corresponding displacement of the rotation vector in space (that is, the arc element of the herpolhode curve) and by C the moment of inertia (equal for all axes) of the spherical top, then evidently

$$di = C\,ds.$$

We then decompose ds into two components parallel and perpendicular to the direction of the instantaneous rotation axis, which may be called ds' and $d\sigma$, respectively. We then have, according to the definitions of the acceleration resistance and the deviation resistance,

$$H = C\,\frac{ds'}{dt}, \qquad K = C\,\frac{d\sigma}{dt}.$$

Now ds' is nothing other than the change in the rotational velocity Ω during the time dt. Thus we obtain for the *magnitude of the acceleration resistance* the formula

$$H = C\,\frac{d\Omega}{dt} \tag{1'}$$

or also

$$H = \frac{d|i|}{dt}. \tag{1''}$$

The *axis* of the acceleration resistance coincides with the axis of the instantaneous rotational velocity; the *sense* is opposed to the sense of the instantaneous rotational acceleration. *The acceleration resistance thus always appears when we accelerate the rotation of the spherical top.* A change in direction of the rotation vector with unchanging length of this vector has no influence on the acceleration resistance. This remark serves as an explanation of the chosen designation.

More important for us is the other component of W, the *deviation resistance.*

To calculate the magnitude K of the deviation resistance, we remark that we can conceive $d\sigma$ as the arc element of the herpolhode curve for the *uniform* precession that has the same direction and magnitude of the rotation axis as our *uniformly accelerating* precession at the

considered moment. If we denote the constants of this precession by μ, ν, ϑ, then, according to equations (3) and (5) of page 50,

$$d\sigma = \sqrt{(d\pi)^2 + (d\kappa)^2 + (d\varrho)^2} = \mu \sin\vartheta \, d\psi = \mu\nu \sin\vartheta \, dt.$$

Thus we have for the *magnitude of the deviation resistance* of the spherical top the simple formula

$$K = |C\mu\nu \sin\vartheta|. \tag{2'}$$

For what concerns the *axis of the deviation resistance*, there first follows, from its definition, that this axis stands perpendicular to the axis of the instantaneous axis of rotation; on the other hand, it follows from the preceding consideration that the axis of the deviation resistance falls in the tangent plane to the herpolhode cone of the uniformly accelerated precession lying through the instantaneous rotation axis, or, which is the same, to the herpolhode cone of the original motion. We have stated a completely analogous characteristic previously for the centrifugal force, whose direction, as mentioned, lies in the osculating plane to the trajectory; that is, in the plane passing through the velocity directions at two neighboring times. The previous osculating plane corresponds in the present case, however, to the cited tangent plane, in so far as it passes through two neighboring rotation axes.

In the last section of this chapter, we will enter into still more detail concerning the characteristic experimental consequences that result from the direction of the axis of our deviation resistance.

Finally, we will establish the *sense of the deviation resistance* through a simply applied geometric rule. The sense is naturally determined, originally, through equation (1′); that is, it is opposite to the component of di perpendicular to the rotation axis. But we can also speak, instead, in the following manner: *the deviation resistance always acts in such a sense that it presses the polhode cone to the herpolhode cone.*

In fact, we consider an arc element of the herpolhode curve for our enforced motion. We will choose our standpoint in space so that the direction in which our arc element runs from the endpoint of the rotation vector is oriented toward us. Since the rotation vector is always drawn toward the side from which the rotation appears to occur in the clockwise sense, the polhode cone necessarily lies, for our choice of the standpoint, to the right of the herpolhode cone. Now, the change in the impulse vector and the change in the rotation vector are proportional

for the spherical top. Our resistance, which is opposed in sense to the impulse change, thus acts, as seen from our standpoint, in the counterclockwise sense. It seeks, therefore, to turn the right-hand lying polhode cone to the left, and thus presses the polhode cone, as claimed, to the herpolhode cone.

Since the calculation of the deviation resistance is concerned merely with the components of the impulse change perpendicular to the rotation axis, a simultaneous acceleration of the rotational velocity has absolutely no influence on the magnitude of our resistance. Rather, our resistance appears only for such motions that are associated with a change of direction of the rotation axis. *We must therefore seek the origin of the deviation resistance, similarly to the centrifugal force, in the change of position of the velocity vector.* We wished to give expression to this fact in the choice of the appellation.

The discussion of this section does not carry over well, as mentioned, to the *generalized top.* We can obviously, to be sure, also decompose the inertial resistance for the generalized top into one component that has the direction of the instantaneous rotation axis, and a second component that stands perpendicular, and thus absorbs no exertion of work. However, the names acceleration resistance and deviation resistance would already be unjustified for these two components. Namely, since the direction of the impulse vector and the rotation vector now fall, in general, apart from one another, the former component would not be exclusively produced by a change of length of the rotation vector, and the latter would not be exclusively produced by a change of direction of the rotation vector. Also, the magnitudes of these two components may not be determined by such simple formulas as for the spherical top. We thus prefer, for the generalized top, to speak simply, as in the previous section, of the *resistance*, and refrain from a further decomposition of this resistance.

§6. The deviation resistance for regular precession of the symmetric top

The discussion of the fourth section contains, in particular, all that there is to say about the enforced motion of the symmetric top. If we enter into the latter in more detail here, this is done only because especially simple results arise for regular precession of the symmetric top, which provide us with a first orientation to the motion of the heavy symmetric top.

We first assume that no external forces act, and give our top an arbitrary regular precession about the vertical with our hand. The impulse vector will thus be led around in a circle about the vertical; its end-point proceeds, at each moment, perpendicularly to the instantaneous rotation axis; its length is not changed. The resistance that results from this impulse change is caused merely by the repositioning of the impulse, and should, correspondingly, again be specially designated as deviation resistance.

The *axis of this deviation resistance* is perpendicular, as was just remarked concerning the direction of the impulse change, to the rotation axis as well as to the vertical; it thus falls, as we may say concisely, in the *direction of the line of nodes.*

The *magnitude and sense of the resistance*, which we denote by the magnitude and sign of the letter K, result from the values of W_X, W_Y, W_Z in equations $(2')$ of page 165. Namely, there follow, since the line of nodes (cf., for example, Fig. 3 of page 18) encloses the angle φ with the X-axis and the angle $\frac{\pi}{2}+\varphi$ with the Y-axis,

$$K = W_X \cos\varphi - W_Y \sin\varphi.$$

Since regular precession is treated, we must insert here into the expressions for W_X and W_Y the values for p, q, r from equations (4) of page 50, and for L, M, N the corresponding values Ap, Aq, Cr. We thus obtain, after a small calculation,

$$K = -C\mu\nu\sin\vartheta - (C-A)\nu^2\sin\vartheta\cos\vartheta. \tag{1}$$

One notes that the meaning of the letter K is somewhat changed with respect to the previous section, in so far as we now attach to K a specified sign. A positive value of K signifies that our deviation resistance acts in the clockwise sense as seen from the half-ray of the line of nodes, and a negative value signifies that it acts in the counterclockwise sense. The resistance determined in this manner will be felt as a pressure on our hand in the guidance of the force-free top.

In the special case $C = A$, our equation (1) naturally reduces to equation $(2')$ derived for the spherical top on page 173, in that the second term on the right-hand side of (1) then vanishes. This second term thus arises from the deviation of the ellipsoid of inertia from spherical form, and should be designated, for example, as the "*ellipsoidal component of the deviation resistance.*" The first term, which signifies

the value of the resistance for the spherical top, should, correspondingly, be called the "*spherical component*."

There are still two other cases in which the deviation resistance reduces to its spherical component. This obviously occurs, first, when $\vartheta = \frac{\pi}{2}$, so that the figure axis is perpendicular to the axis of the regular precession. But, further, a similar result obtains if the rotation component μ predominates extraordinarily in relation to all other quantities found in our formula. This latter case is actually the rule in applications, since one generates, by means of the commonly used driving mechanisms, an impulse vector of very large magnitude, which falls nearly in the direction of the figure axis.

We return here once again to *natural precession*, which the force-free symmetric top executes without exterior forcing. The condition for this motion is obviously (one compares the development on page 167) $K = 0$, or

$$0 = (C\mu + (C - A)\nu \cos\vartheta)\nu \sin\vartheta.$$

Of all possible precessional motions, an entirely determined class is separated by the preceding equation as natural force-free precession. We were already led to this same condition in the second section (see eqn. (11) of page 153).

We now consider a *heavy symmetric top*, and impart to this a regular precession about the vertical with arbitrary constants. From statics (cf. page 87) we know that the gravitational action on the individual mass elements causes a resulting turning-force whose axis, for the symmetric top, is the line of nodes, and whose magnitude and sense are determined by $P \sin\vartheta$. Here $P > 0$ signifies that the center of gravity lies above the support point, and $P < 0$ that the center of gravity lies below the support point.

The resistance that we must overcome for the generation of a regular precession of our heavy top—we can designate it again specially as deviation resistance—again falls in the direction of the line of nodes, since both the turning-moment of gravity and the impulse change have, according to the above, the line of nodes as their axis for regular precession about the vertical. The magnitude and sense of our resistance result now from equation (3′) of the fourth section. If we insert there for D the just given value $P \sin\vartheta$, and for the rate of change of the impulse the quantity $-K$ calculated above, there follows

(2) $$W' = K + P \sin\vartheta.$$

The resistance that we will feel in the guidance of the heavy top is therefore different from that for the same guidance of the force-free top.

We will also make here the transition to the natural motion of the heavy top, and therefore ask for the precessional motion that the heavy top is able to execute without external forcing. The condition for this motion is (in conformity with D'Alembert's principle) $W' = 0$. With consideration of (1) and (2), we can thus write, if we assume that ϑ differs from zero, and that the precession does not, therefore, degenerate into a pure rotation about the vertical,

$$P = C\mu\nu + (C - A)\nu^2 \cos\vartheta. \tag{3}$$

The heavy symmetric top without external forcing may thus execute (for a mass distribution determined by given values of A, C, and P) only a specific class of regular precessional motion; namely, only a precession whose constants μ, ν, ϑ are bound by the relation (3).

In total, this natural precession of the heavy top depends on four arbitrary parameters. These are, for example, the initial values of the angles φ, ψ, ϑ and the angular velocities μ and ν, which latter represent, because of (3), only one arbitrary parameter. We express this fact by saying *there are ∞^4 possible precessional motions for an individual heavy top.*

In contrast, the general natural motion of the heavy top depends on six arbitrary parameters. Namely, since the the top is a system of three degrees of freedom, we can prescribe, in addition to the initial values of φ, ψ, ϑ, their initial velocities arbitrarily, from which the further course of the motion is determined through the differential equations of mechanics. *There are thus ∞^6 possible natural motions for the heavy top.*

We thus conclude that regular precession is only a *special form of motion of the heavy top*, a *particular solution of its differential equations.*

If we make this same enumeration for the symmetric gravity-free top, then we must consider, in addition to the initial values of the angles φ, ψ, ϑ and the angular velocities μ and ν bound through the relation $K = 0$, that the axis of precession can have any desired position through O (while this axis must necessarily coincide with the vertical for the heavy top because of the direction of the action of gravity). This

circumstance increases the number of arbitrary constants for regular precession of the force-free top to 6. *There are thus* ∞^6 *natural precession motions.* We understand completely well, on this basis, why regular precession can be the most general form of motion of the gravity-free symmetric top.

Regular precession will later become important for us as the simplest, although particular, motion of the heavy top. We will thus enter into the possible values of the precession constants in detail, in which we let these constants arise from the precession constants for the force-free top by a continuous transition.

We will thus imagine μ and ϑ as fixed, and ask how the value of ν is changed by the introduction of gravity. We will assume, moreover, the value of μ to be positive.

If we first take $P = 0$, then equation (3), quadratic in ν, yields two values of this constant; namely,

$$\text{a)}\ \ \nu = 0, \qquad \text{b)}\ \ \nu = \frac{C\mu}{(A-C)\cos\vartheta}.$$

The value a) corresponds to a simple rotation of the top about the figure axis, and the value b) corresponds to an actual precessional motion. We next assume P different from zero. We then find as roots of (3) the two values

$$\nu = \frac{C\mu \mp \sqrt{C^2\mu^2 - 4P(A-C)\cos\vartheta}}{2(A-C)\cos\vartheta}. \tag{4}$$

There are thus (for a given mass distribution and given values of μ *and* ϑ*) either two cases of regular precession or none, according to whether the square root in the preceding expression is real or imaginary.* We will distinguish the two real cases, according to the value of ν, as "*slow*" and "*fast precession.*" For sufficiently small values of P, the slow precession is that which develops by continuous transition from the simple rotation of the gravity-free top about the figure axis, whose precession constant therefore corresponds to the root designated a) above.

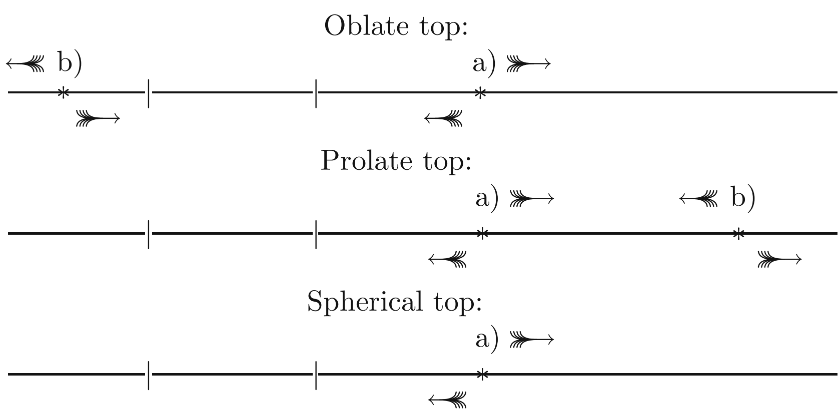
The case of the spherical top ($A = C$) requires a special consideration. In this case, equation (3) becomes linear in ν and yields the value

$$\nu = \frac{P}{C\mu}. \tag{5}$$

The comparison with the oblate or prolate top allows us, however, to speak also in this case of a second root, which, however, has become infinitely large ($\nu = \pm\infty$). We will be able to denote the precessional motion corresponding to the first root as slow precession, since it is transformed for $P = 0$ into the simple rotation a); the root $\nu = \pm\infty$ represents, in every case, a fast precession.

The following schema, whose validity is easily verified from equations (4) and (5), depicts the positions of the roots a) and b) for $P = 0$, as well as (through the sense of the attached arrows) the directions in which a) and b) move with the appearance of gravity in the cases of the oblate, prolate, and spherical top. The arrows above and below refer to the directions in the cases $P > 0$ and $P < 0$, respectively.

Peri-	Anti-	Hypo-	Epi-cycloidal motion



In connection with these results, we remark that the values of $\dfrac{\nu}{\mu}$, which for the force-free top were restricted to the domains of epi- and pericycloidal motion (cf. page 153), can encroach upon the regions of anti- and hypocycloidal motion in the case of the heavy top, and that, therefore, all kinematically possible precessional motions can be realized kinetically by the heavy top.

It can, perhaps, be initially astonishing that the force of gravity acting entirely in the vertical direction can maintain a regular precession of the top about the vertical; that is, a motion for which the individual mass elements progress, on average, in the horizontal direction. However, this state of affairs is in no way more remarkable than the

well-known fact in particle mechanics that a mass particle which moves under the influence of a central force orbits, for the appropriate initial impulse, in a circle about the center of attraction, and can therefore continuously progress in a direction exactly perpendicular to the direction of the external force. Just as the possibility of this motion is given by the condition that the centrifugal force must always be in equilibrium with the attraction force, so the condition for which we can realize regular precession as a possible motion of the heavy top is that the inertial resistance for the regular precession should always maintain equilibrium with the applied turning-moment about a horizontal axis caused by the force of gravity. It is to be remarked that neither the circular motion of the particle nor the regular precession of the top represents the most general motion of the particle or the top, but rather that the two motions occur only for an appropriate choice of the initial state.

In conclusion, we may consider the regular precession of the heavy top from the point of view of the so-called stability of the rotation axis discussed in the preceding chapter. According to the above, a regular precession ($\nu \gtrless 0$) can develop from a simple rotation ($\nu = 0$) with the appearance of gravity, in which the rotation axis will gradually be led around a circular cone about the vertical. We therefore see, in agreement with a remark on page 136, that a continuously applied force can, in the course of time, occasion a complete repositioning of the rotation axis, as large as the original rotational velocity μ or as small as the external influence P may be. *It is therefore certainly not possible to speak of a permanence of the rotation axis with the appearance of a continuous external disturbance.*

§7. A new derivation of the deviation resistance for regular precession of the symmetric top. The Coriolis force

Since the concept of the deviation resistance for the regular precession of the symmetric top will play a not inconsiderable role in the following, we wish to make clear its origin in yet another manner, in which we imagine the top resolved into its individual mass particles, and ask for the forces with which these particles resist the enforced motion. We will limit ourselves, for the sake of simplicity, to the special case $\vartheta = 90°$ of regular precession, in which, according to page 176, the deviation resistance reduces to its spherical component.

We first consider an individual mass particle P, and force this particle to traverse a path composed from two circular motions. The point should progress in a vertical plane E with constant velocity in a circle of radius R and midpoint O; at the same time, the plane E should turn with uniform velocity about a vertical axis A through O.

We wish to calculate the resistance with which the particle resists this composed motion. We must, in any case, find as components of this resistance the two centrifugal forces that correspond to the component motions of P; namely, the circulation in the plane E about O, and the circulation in space occurring about A. But it appears that an additional force results from the coexistence of the two components of the motion, a force that is called, after its discoverer, the *Coriolis force* or the *composed centrifugal force (force centrifuge composée)*.

We specify the position of the point P in the plane E by the angle φ that its binding line with O makes with respect to the initial position of this line. We wish to assume that the binding line OP is initially perpendicular to the axis A. If we denote by μ the constant angular velocity with which P circulates in the plane E, then we have

$$\varphi = \mu t.$$

We determine, on the other hand, the position of the plane E in space by the angle ψ through which E has turned from the initial position E_0. If we denote by ν the constant angular velocity with which E turns about A, then we have

$$\psi = \nu t.$$

We may arrive at the calculation of the Coriolis force in the shortest way by the following small coordinate calculation. We will give a direct geometric derivation at the conclusion of the section (cf. pages 186 and 187).

We take the rotation axis A as the z-axis of a spatial coordinate frame whose origin lies at O, and whose xz-plane coincides with E_0. The coordinates of P in this frame evidently become

$$\text{(1)} \qquad \begin{cases} x = R\cos\varphi\cos\psi, \\ y = R\cos\varphi\sin\psi, \\ z = R\sin\varphi. \end{cases}$$

The impulse i of the particle for our composed motion then has, if we denote the mass of the particle by m, the components

$$\begin{aligned}
[X] &= mx' = -\, m\mu R\sin\varphi\cos\psi - m\nu R\cos\varphi\sin\psi,\\
[Y] &= my' = -\, m\mu R\sin\varphi\sin\psi + m\nu R\cos\varphi\cos\psi,\\
[Z] &= mz' = \quad m\mu R\cos\varphi.
\end{aligned}$$

The resistance W with which the point opposes the enforced motion is determined, as earlier, through the vector equation

$$W = -\frac{di}{dt};$$

its components therefore become

$$(2)\quad \begin{cases}
W_x = -\dfrac{d[X]}{dt} = m\mu^2 R\cos\varphi\cos\psi - 2m\mu\nu R\sin\varphi\sin\psi\\
\qquad\qquad\qquad\qquad + m\nu^2 R\cos\varphi\cos\psi,\\
W_y = -\dfrac{d[Y]}{dt} = m\mu^2 R\cos\varphi\sin\psi + 2m\mu\nu R\sin\varphi\cos\psi\\
\qquad\qquad\qquad\qquad + m\nu^2 R\cos\varphi\sin\psi,\\
W_z = -\dfrac{d[Z]}{dt} = m\mu^2 R\sin\varphi.
\end{cases}$$

The terms in these expressions with the factor μ^2 obviously correspond to the centrifugal force that would arise from the circulation of P in the plane E, in so far as we imagine the latter standing still in space. In fact, these terms are the components of a force of magnitude

$$(3)\qquad m\mu^2 R,$$

whose direction is specified with respect to the coordinate axes through the direction cosines

$$(3')\qquad a = \cos\varphi\cos\psi,\quad b = \cos\varphi\sin\psi,\quad c = \sin\varphi.$$

These, according to equations (1), are directly the direction cosines of the binding line OP.

On the other hand, the terms with the factor ν^2 correspond to the centrifugal force that would result from the circulation of P about the axis A, in so far as we regard the point P as stationary in the plane E. In fact, these terms give a force of magnitude

$$(4)\qquad m\nu^2 R_1,$$

where $R_1 = R\cos\varphi$ signifies the distance of the point P from the axis A. Its direction cosines with respect to the coordinate axes are

$$(4')\qquad a' = \cos\psi,\quad b' = \sin\psi,\quad c' = 0;$$

that is, those of a line passing through P perpendicular to the axis A.

There yet remain the middle terms with the factor $\mu\nu$, which owe their origin to the fact that the plane E is not fixed in space, nor is the point P fixed in the plane E. If we compose the respective components, there results a force of magnitude

$$(5) \qquad -2m\mu\nu R \sin\varphi,$$

whose direction cosines with respect to the coordinate axes are

$$(5') \qquad a'' = \sin\psi, \quad b'' = -\cos\psi, \quad c'' = 0.$$

The force is perpendicular to the direction (3′) as well as to the direction (4′), and lies along the normal to E. It is this force that one designates as the *Coriolis force*.

If we ascribe rigidity and a certain thickness to the plane E, and imagine the circle that P describes as a channel engraved in E, then the force (3) will be canceled by the rigidity of E, and the force (4) will be canceled in part by this rigidity, and in part by the guidance by means of which P is pushed along the groove uniformly. The force (5), on the other hand, presses perpendicularly to the plane E, and will become sensible if we provide the rotation of this plane about the axis A by hand. This force acts, depending on the quadrant in which P is located instantaneously, to accelerate or retard the rotation of E.

We remark, further, that we learn nothing fundamentally new from the introduction of the Coriolis force. Just as the introduction of the centrifugal force signifies nothing that we do not know already from the theory of the impulse, so we learn nothing from the consideration of the Coriolis force that is not included in the general theory of centrifugal forces. The necessity of the Coriolis term follows simply from the fact that the centrifugal force, which corresponds to the actual motion of the particle in space, is not simply equal to the resultant of the two centrifugal forces that result if we let, on the one hand, P rotate in the stationary plane E, and, on the other hand, let E rotate with a fixed position of P.

The independent introduction of the Coriolis force is generally pronounced only from a special point of view that one can adopt with respect to the composed motion described above, and which we wish to designate as the *standpoint of relative motion.* One may, to remain with the above example, treat of the motion of the particle P in the plane E,

so that one does not step outside the plane E in the consideration. If this plane were at rest, one would have to take account of the first of the above centrifugal forces. But because of the rotation of the plane E, a correction in the form of the Coriolis terms must be added, which owe their origin to the composition of the two motions. The third term, which is of order ν^2, can be neglected in relation to the first two under certain circumstances; namely, whenever the motion of E in space proceeds relatively slowly, and when ν is therefore sufficiently small in relation to μ.

Moreover, the Coriolis approach*) goes much further than our presentation of it. While we considered an individual mass particle that executes a prescribed circular motion in a uniformly rotating plane, the Coriolis approach applies to an arbitrary system of mass particles that describe an arbitrary motion with respect to a coordinate system that moves, in its turn, arbitrarily. —

The class of relative motion obviously includes *any motion on the surface of our Earth.* One always treats of such a motion as if the surface of the Earth were at rest, and includes the influence of the rotation of the Earth into the calculation afterward by the inclusion of the Coriolis force and the centrifugal force of the Earth's rotation. If, in particular, the moving body is forced to progress uniformly on a meridian, we have exactly the phenomena assumed above.

We now come to the application of the preceding consideration to the theory of the top. For this it is permissible to choose a *geographic garb*, and retain the just mentioned representation of a motion proceeding on the surface of the Earth.

We will imagine that the Earth is surrounded along a meridian by a channel with an everywhere equal rectangular cross section, in which we let water circulate with uniform velocity, while at the same time the Earth rotates about its axis with uniform velocity. We denote the center of the Earth, which we imagine as fixed in space, by O. Let the velocity of the rotation of the Earth be ν; the angular velocity with which the water flows in our channel as seen from O is μ. The Earth

*) Cf. his treatise in the 24th volume of the Journal de l'École Polytechnique of the year 1835: *Sur les équations du mouvement relatif des systèmes de corps.*[79]

rotates in the counterclockwise sense as seen from the North Pole, as is indicated in the adjacent figure by an arrow. The water flows, in the forward half of the Earth shown in the figure, from south to north. The designations "right" and "left banks" should be chosen for our channel, as is usual in geography, from the point of view of an observer looking in the flow direction. The plane of our meridian will be denoted concisely by E.

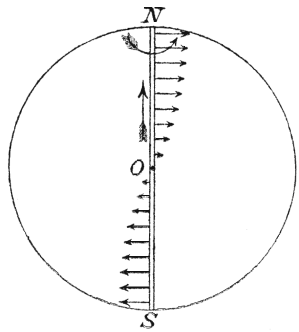


Fig. 21.

We will, moreover, so as not to have to enter into hydrodynamic questions and to be able to apply the preceding considerations directly, assume our channel to be infinitely thin, and treat of our water flow as if it consisted of a series of individual successive discrete mass elements.

Since there is no reason why one water element in our channel should overtake another, the spacing of the individual elements remains constant in the motion. The water, we can say, pushes itself forward in its channel exactly as though it were a rigid body. As a result, the forces that act on an individual water element will be the same as for a rigid body of a corresponding mass distribution that turns about the fixed point O; that is, as for a top. And indeed, the mass of our top, which corresponds in this sense to the water in our channel, is to be imagined as concentrated on a circle that has as its radius the Earth radius R and as its center the center of the Earth O. Since we have, moreover, to imagine the mass uniformly distributed on the given circle, our top must be designated, in particular, as a symmetric top; its figure axis is the line perpendicular to the meridian plane E. The motion of the top, which corresponds to the composed motion of the water circulation and the rotation of the Earth, thus represents a simple regular precession, in which the figure axis turns about the Earth axis with velocity ν and perpendicular inclination to this axis, while at the same time the top rotates about the figure axis with the velocity μ.

We return again to the geographic image, and ask for the *force with which the water presses on the lateral walls of the channel.*

The total force which an individual water element P exerts in opposition to the assumed motion (that is, the previously calculated resistance W) consists of three parts. We have, first, a centrifugal force that arises from the circulation of the water alone and acts in the direction OP, and, second, a centrifugal force that is caused by the rotation of the Earth alone and is directed perpendicularly to the axis of the Earth. These forces obviously give no lateral pressure on the walls of the channel, at least in so far as we represent, as agreed above, the water flow as a simple series of discrete mass elements. They are manifested merely in an apparent decrease in weight of the water, or, what signifies the same thing, in a pressure against the upper wall of the channel. (For the associated motion of the top, this pressure would be directly resisted by the rigidity of the material.)

But, in addition, we have yet the Coriolis force to consider. According to (5), this amounts to

$$-2\mu\nu R\sin\varphi\, dm, \tag{6}$$

where φ denotes the geographic latitude and dm denotes the mass of P. It is directed, according to the above, perpendicularly to the meridian plane E, and therefore perpendicularly directed with respect to the lateral wall of the pipe, and changes its sense as $\sin\varphi$ changes its sign. *In the Northern Hemisphere* it presses throughout *against the right bank, and in the Southern Hemisphere against the left.* The magnitude and sense of this force in the different positions of our channel are indicated in the preceding figure by arrows.

We can still further illustrate the origin of the force (6) in our geographic image through the following deliberation, which we divide into 2 steps.

1) A water element that proceeds along our rotating Earth meridian has, at each location, one velocity component that falls in the direction of the meridian, and a second component that falls in the direction of the parallel circle. These two components are obviously equal, respectively, to the product of the angular velocity μ and the radius of the meridian, and the product of the angular velocity ν and the radius of the parallel circle. The corresponding components of the impulse will thus be given by

$$R\mu\, dm \quad \text{and} \quad R\nu\cos\varphi\, dm.$$

While the element starting from the equator wanders towards the North Pole, the impulse component falling in the direction of the parallel circle

will be successively changed, in that the distance of our element from the axis of the Earth will be successively diminished. The rate of change of this velocity component obviously amounts to

$$\frac{d}{dt} R\nu \cos\varphi \, dm = \mu \frac{d}{d\varphi} R\nu \cos\varphi \, dm = -\mu\nu R \sin\varphi \, dm.$$

The element has, as we can say, an excess in lateral velocity as a result of the rotation of the Earth, which must be reduced in the progression along the channel. The corresponding excess in the impulse signifies as much as an infinitesimal impact that the element exerts against the channel walls; the corresponding rate of change of the impulse is therefore a continuously acting force perpendicular to the lateral walls.

2) The force thus found represents, as we see, only half of the Coriolis force given above. There is, in fact, still another circumstance to consider, which has, in consequence, an equally large change in the impulse. It is, namely, that at each moment (except for the points of the equator) the *direction* of our meridional channel changes because of the rotation of the earth. If we denote the change in direction occurring in the time dt by $d\chi$, then there results in this time a velocity change again falling in the direction of the parallel circle, which is equal in magnitude to the product of $d\chi$ and the meridional velocity component $R\mu$. Through an elementary geometric consideration, one easily finds for $d\chi$ the value

$$d\chi = -\nu \sin\varphi \, dt.$$

The corresponding rate of change of the impulse, which manifests itself as a lateral pressure, will be

$$R\mu \frac{d\chi}{dt} \, dm = -\mu\nu R \sin\varphi \, dm.$$

The circumstances cited in 1) and 2) together yield, in fact, the force given in (6). We can directly regard the preceding consideration as the previously mentioned new geometric derivation of the Coriolis force.

We now compose the calculated forces on the individual water elements into a resulting turning-force that is applied, in total, to the Earth. We can disregard the usual centrifugal forces, since each such force is compensated by an equal and oppositely directed centrifugal force in the consideration of the total effect on the Earth. The effect depends, therefore, only on the "composed centrifugal forces" given in (6).

To begin, the individual force (6) corresponds, with respect to the assumed fixed center of the Earth, to a turning-force of magnitude

$$-2\mu\nu R^2 \sin\varphi\, dm,$$

which, as one easily reads from the figure, acts in the positive sense about the half-ray $\varphi - \frac{\pi}{2}$. We will decompose this turning-force into two components, in that we project, on the one hand, onto the axis of the Earth, and, on the other hand, onto the "line of nodes"; that is, onto the equatorial line $\varphi = 0$ perpendicular to the axis of the Earth. These two components then become, respectively,

$$2\mu\nu R^2 \sin\varphi \cos\varphi\, dm$$

and

$$-2\mu\nu R^2 \sin^2\varphi\, dm.$$

If we make the same decomposition for each water element dm and compose the resulting turning-forces, then we obtain the desired resultant turning-force. This resultant has, understanding by $2\pi m$ the total mass of the flowing water, the component

$$+2\mu\nu R^2 \int_{-\pi}^{+\pi} \sin\varphi \cos\varphi\, dm = +2\mu\nu m R^2 \int_{-\pi}^{+\pi} \sin\varphi \cos\varphi\, d\varphi = 0$$

in the direction of the axis of the Earth; its second component in the direction of the line of nodes, in contrast, is

$$-2\mu\nu R^2 \int_{-\pi}^{+\pi} \sin^2\varphi\, dm = -2\mu\nu m R^2 \int_{-\pi}^{+\pi} \sin^2\varphi\, d\varphi = -2\pi m R^2 \mu\nu.$$

If we consider, further, that

$$2\pi m R^2 = C$$

signifies the moment of inertia of our water flow with respect to the Earth diameter perpendicular to the meridian plane, then we can say:

The resulting turning-force that is applied to the Earth as a consequence of the assumed water motion is, in total,

$$K = -C\mu\nu,$$

and has the just cited line of nodes as its axis.

If we think in terms of the regular precession of the top that is analogous to our water motion, then we can designate this turning-force K directly as the *deviation resistance* of our water mass. The preceding

formula is obviously in complete conformity with equation (1) of the previous section, if only we take, in that equation, the corresponding condition $\vartheta = \frac{\pi}{2}$ assumed here.

If we compare again the present derivation of the deviation resistance with the previous, we will find that the two are not as different as they may first appear. Here, we started from the impulse of the individual mass element, from which the Coriolis force resulted by differentiation with respect to time. An integration over the total mass of the top was further necessary to find the total resistance of the top. Earlier, we operated with the total impulse of the top, from which the resistance directly followed by differentiation with respect to time. The total impulse, however, was, in its turn, derived (in the previous chapter) from the impulses of the individual mass particle through integration over the mass of the top. *The difference between our present and earlier derivations therefore consists*, disregarding the greater generality of the latter, *fundamentally in a mere interchange of the order of differentiation and integration.*

We must emphasize that the water flow considered here on the uniformly rotating surface of the Earth belongs completely to the category of enforced motion, and that it will be possible, strictly speaking, only through an exterior compulsion. If, namely, the rotation axis of the Earth should always remain in the fixed north–south direction, then the deviation resistance of the water flow must be overcome from the exterior at each moment. Since such a compulsion is in reality not at hand, there would immediately result from the circulation of the water a *deflection of the north–south direction* and a *change in the local geographic latitude.* We will return to the questions associated with this phenomenon later in the astronomical applications of the theory of the top.

The pressure that, according to the above, meridionally flowing water must apply against the lateral walls of a channel is well known in geography, and has given occasion for the statement of the so-called *Beer's Law.*[80] This law has been used to derive — whether correctly or incorrectly, we need not decide here — changes in the course of rivers. The corresponding consideration would hold for a *railroad train**) that travels with constant velocity from the equator toward one of the two poles. Such a train must, if only its velocity is large enough, be thrown

*) Numerical details concerning this are found in A. Ritter, Höhere Mechanik, 1. Teil, p. 153 and ff.[81]

over in the *Northern* Hemisphere toward the *right* as calculated from the direction of travel, and in the *Southern* Hemisphere to the *left*, at each specified point of the surface of the Earth. If we imagine, instead of flowing water, a meridionally circulating flow of air, a similar consideration obtains. The pressure that the channel walls must bear would now be expressed in a deflection of the air stream from its original direction, which is determined in detail by the well known *Buys-Ballot law*.[82]

These and similar things are discussed, in detail, in an interesting and popularly received work*) of Jouffret.

§8. Experimental demonstration of the deviation resistance. The top with one and two degrees of freedom

We wish not only to understand the appearance of the deviation resistance theoretically, but also to verify it practically.

Our purpose here is served by a simple *hand-top*, illustrated in the adjacent figure. It consists of a rotor (visible in the figure in cross section), which is pivoted without friction, as we will assume, in an outer ring. A handle is provided on the ring.

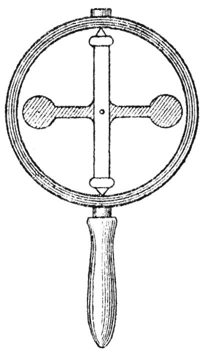


Fig 22.

If we imagine the handle as fixed, then we have before us a *top of only one degree of freedom.* But if we move the handle in an arbitrary way about a point on its axis, as will be done in the following, then we realize the case, already discussed in the fourth section, of a *top with only one natural motion degree of freedom and two enforced motion degrees of freedom.* In fact, we are no longer able (because of the assumed absence of friction in the pivot) to govern the angular velocity about the figure axis by the motion of

*) *Théorie élémentaire du mouvement du gyroscope* etc. Revue d'Artillerie, Bd. 4, 1874. We remark, however, that the Jouffret formulas are not generally reliable. In the calculation of the lateral pressure, only the first of the circumstances mentioned on pages 186 and 187 is considered, and thus the magnitude of the deviation resistance is found too small by a factor of $\frac{1}{2}$.[83]

Cf. also H. Scheffler: *Imaginäre Arbeit, eine Wirkung der Centrifugal- und Gyrokraft.* Leipzig 1866. The "gyroscopic force" of Mr. Scheffler coincides, in essence, with our "deviation resistance." Mr. Scheffler characteristically designates the effort that is necessary for the overcoming of the deviation resistance, since it signifies no work in the usual sense of the word, as "imaginary work."[84]

the handle after we have set the rotor into rotation.

We wish to assume that the impulse originally imparted to the top is quite considerable, and, in any case, much larger than the impulse changes that we will cause by the subsequent enforced motion. If we now swivel the apparatus rapidly back and forth by hand, then we feel a strong pull on our hand, from which we recognize that the top exerts a resistance to the motion enforced on it. This resistance obviously belongs to the category of the forces considered in the previous section, and has its basis in the changes of the rotational impulse enforced by the external motion. It may not be attributed, for example, to the usual inertial resistance of nonrotating bodies, with which it already does not coincide in direction, and which it greatly exceeds in magnitude.

In particular, we can imitate *regular precession* in the way that we lead the top around our body with an outstretched arm. Here also we feel a pull, which we can pronounce in particular, according to §6, as deviation resistance. The arm is pulled slightly upward or downward if we lead the top in the one or the other sense about the vertical. However, this pull will be distinctly sensible for normal dimensions of the apparatus only in the case of a very rapid rotation.

The *sense of the pull* is particularly conspicuous here. If we lead the handle in a horizontal direction, the pull acts in the vertical direction. *It appears that the top always wishes to evade the instantaneous direction of the motion perpendicularly.* This apparently paradoxical behavior, which has already confronted us in a similar manner many times, is explained by the very concept of the deviation resistance. In fact, the impulse change for our regular precession occurs in the horizontal direction; the turning-force of the deviation resistance thus acts about a horizontal axis and seeks to raise or lower the handle of the top, according to whether the precession is progressive or retrograde.[85]

The following experimental arrangement is much more sensitive. We lay the top with its handle on two fingertips of the right hand and the fixed ring on two fingertips of the left hand. If we now rotate the axis of the instrument, in that we displace the fingertips oppositely to each other in the horizontal plane, then we sense a distinct pressure on one

or the other hand. We can conceive the lateral motion of the fingertips as the beginning of a regular precession in which the figure axis is perpendicular to the axis of the precession. The pressure signifies to us the existence of the deviation resistance.

The places of the fingertips can also be taken by the two pans of a scale. We lay the hand-top, after we have transferred to it a strong rotation, with its handle on one pan of the scale and its ring on the other. We now turn the table on which the scale stands. One pan of the scale will immediately begin to sink slightly.[86] —

We next treat of the top represented in the adjacent figure. We will regard the bearing of the rotor (shown in cross section) and the bearing of the inner ring as frictionless, and regard exterior forces as not present. Concerning the mass distribution, we make, in order to be able to treat of our apparatus as a single rigid body, the assumption already discussed in the Introduction (cf. page 3), that the mass of the rotor considerably predominates the mass of the inner ring.

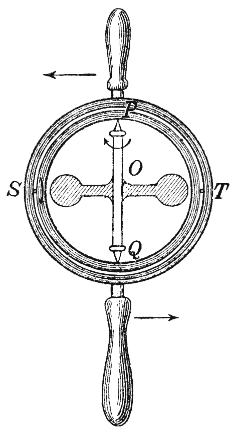


Fig. 23.

If we imagine the handle to which our top is attached as fixed in space, then we have a *top with two degrees of freedom.* But if we turn, as will later be done, the handle itself in the plane of the outer ring, then there is added to the *two natural motion degrees of freedom* an *enforced motion degree of freedom.* The interest that the apparatus may claim, and which justifies its more precise study in this place, is that Lord Kelvin seeks to support and make plausible his ingenious "kinematic theory of matter," which we will consider in a later chapter,[87] by means of the simple model just described.

We denote the figure axis of the rotor by PQ, and the rotation axis of the inner ring by ST. The two lines intersect in the midpoint O of the top. Let the moment of inertia of the rotor with respect to the axis PQ be equal to C, and the moment of inertia of the rotor with respect to the axis ST be equal to A. The inner and outer rings originally lie, as in the figure, in a plane.

We first wish to make clear *that the natural motion of our top with*

two degrees of freedom (for a fixed position of the handle) is a regular precession, in which the rotor turns with a uniform velocity about the axis PQ, and the inner ring rotates with an independent uniform velocity about the axis ST. We take the criterion for the natural motion of our top from §4. While for the natural motion of the force-free top with *three* degrees of freedom the change of the impulse must simply be equal to zero, for the natural motion of our top with *two* degrees of freedom only those components of the impulse change that correspond to a possible motion must vanish, while an impulse change that would correspond to a rotation of the attached handle is simply noticeable as a pressure on the handle. If we would use the manner of expression of *D'Alembert's principle* (cf. page 167), we could also formulate this criterion as follows: the system of the inertial resistances of all the mass particles must maintain equilibrium with the system of external forces; that is, in our case, with the mechanism providing the fixed position of the handle.

In our regular precession, the impulse will now be led on a circular cone around the axis ST, and continuously lies in a plane containing the axes ST and PQ. As a consequence, the impulse change is continuously perpendicular to these two axes. The axis of rotation that corresponds to this impulse change coincides with the axis of the impulse change itself; it is, therefore, likewise perpendicular to the two possible rotation axes PQ and ST. Our criterion above is consequently fulfilled, and our claim is proven.

We can thus say: if we have imparted to the rotor arbitrary angular velocities μ and ν about the axes PQ and ST, then these angular velocities always retain their original values. The negatively taken rate of change of the impulse, that is, the *deviation resistance of the regular precession*, must, therefore, always be overcome by the mechanism providing the fixed position of the handle. The magnitude of this resistance, in the present case $\vartheta = \frac{\pi}{2}$, is (cf. page 175)

$$|K| = |C\mu\nu|. \tag{1}$$

Its direction, since at the beginning of the motion the planes of the interior and exterior rings coincide, is initially perpendicular to the plane of the outer ring. Its inclination with respect to this plane will change, in contrast, as soon as the figure axis steps outside the plane

of the exterior ring in the course of the precession. To give, in addition, the sense of the resistance, we assume that the angular velocities μ and ν occur in the clockwise sense as seen from P and S. The impulse change at the beginning of the motion is then directed perpendicularly from the plane of the diagram outward to the viewer. The deviation resistance, which is opposed to the impulse change, is thus applied about the forward normal to the plane of the diagram in the counterclockwise sense, and initially seeks to move the handle as indicated by the horizontal arrows of the figure.

To simplify the manner of expression, we will agree to consider only the beginning of the motion; that is, only that time in which the figure axis remains near the plane of the exterior ring. Under this assumption, we can now say concisely that our resistance is perpendicular to the plane of the outer ring, and seeks to turn the handle in this plane in the counterclockwise sense as seen from the front.

We can here again recall the alleged invariability of the rotation axis. For the regular precession of our top with two degrees of freedom, the instantaneous rotation axis (at least for small values of $\nu:\mu$) will be led about the horizontal axis ST on a very flat circular cone. The external influence that is necessary here consists in the continuous overcoming of the deviation resistance. *We can therefore effect, simply by fixing the handle, that is, without performing the least amount of work, that the rotation axis will be transformed from its original position into a nearly opposite position.*

We will now assume, conversely, that no rotation about the axis ST is present at the beginning of the motion, but that the top is given a strong rotation μ about the axis PQ, which occurs in the clockwise sense as seen from P. We imagine the locking of the handle as lost, for an instant, and apply to the handle an enforced motion of the simplest type. Namely, we suddenly turn the handle in the plane of the outer ring through a small angle v in the clockwise sense as seen from the front, and then immediately fix the handle again.

The impulse of the top is naturally composed, according to the parallelogram law, with the turning-impacts that correspond to our motion v. Now, there corresponds to the generation of the motion v a certain turning-impact, and to the immediately following cessation of this motion the equal and oppositely directed turning-impact. In total,

therefore, the impulse vector retains its original position in space. But at the same time, its position with respect to the top will be changed: while the impulse vector originally fell, by assumption, in the direction of the figure axis, it now forms the angle v with this axis. Thus the projection of the impulse vector on the rotated direction of the figure axis has, for sufficiently small v, approximately the magnitude of the original impulse $C\mu$. The component of the impulse falling in the direction of ST, perpendicular to the figure axis, has, in contrast, the magnitude $C\mu v$.

To the latter impulse component there corresponds a rotational velocity ν, which occurs about the axis ST in the clockwise sense as seen from S, and which is determined by the equation

$$A\nu = C\mu v. \tag{2}$$

The inner ring begins to move with this velocity ν after we have turned the outer ring. The rotor continues, in contrast, to rotate about its figure axis with angular velocity μ, since the impulse component in the direction of the rotated figure axis still amounts to $C\mu$. If we no longer change the position of the handle in space, then μ and ν retain the given values. There appears, therefore, a regular precession of exactly the nature considered above.[88]

We remark, further, that the velocity ν is very small, since we assumed v to be a very small angle. Our agreement above, that we wished to consider only that part of the regular precession for which the figure axis remains near the plane of the outer ring, now implies no essential restriction.

We saw above that every regular precession (μ, ν) of our top, for a fixed position of the handle, produces a deviation resistance, which, under the present circumstances, initially seeks to turn the apparatus in the plane of the outer ring in the counterclockwise sense. The magnitude of this deviation resistance is, in consequence of equations (1) and (2),

$$|K| = \frac{C^2\mu^2}{A} v; \tag{3}$$

it is therefore proportional to the angle v, and the sense of our resistance, as we saw above, is opposed to the original rotation v, and seeks to cancel the latter.

We will later have use for this remarkable property of the top of two degrees of freedom in the representation of the Kelvin elasticity model.

We summarize the property here once more as follows, where we permit ourselves the simplification of the manner of expression that results from our agreement of a limited extent of the observation time:

If we turn our handle with a rotating flywheel through a small angle in the plane of the outer ring, then it resists with a force proportional to the turning angle. We must, to keep the handle in its new position, constantly exert a force. If we discontinue this force, then our handle again seeks its initial position.

Our handle thus has a certain specific capability of resistance against changes of direction, a kind of absolute orientation in space.

Addenda and Supplements[89]

To the Foreword

The statement of Sir John Herschel cited at the conclusion is found in the Outlines of Astronomy, London 1850, Chapter V, No. 317.

A monograph on the top has also appeared recently in England: H. Crabtree, An elementary treatment of the theory of spinning tops and gyroscopic motion, Longmans, Green Co., London 1909.[90] The phenomena of the top are approached here by way of the impulse theorems, as recommended in our Foreword. More systematic and more detailed in its reasoning than the well-known work of Perry (cf. the note on page 134), the book attains to the calculation of the general motion of the top and to the description of some technical applications, such as the ship's gyroscope and the monorail. It is especially rich in examples from daily life, whose essential dynamic traits are emphasized, but naturally makes, due to its greatly limited space and its elementary character, no claim to the completeness that we have sought with wide-ranging analytic means.

To the Introduction

In addition to the textbooks on mechanics cited in the Introduction, a particularly noteworthy and comprehensive book by A. G. Webster has now appeared in which the theory of the top and some of its applications have also found ample space: The dynamics of particles and of rigid, elastic, and fluid bodies, Leipzig 1904. A German edition by C. H. Müller under the title "Lehrbuch der Mechanik als Einführung in die theoretische Physik," in two parts, is in preparation and will appear soon by the same publisher.[91]

The technical questions of mechanics are naturally treated in the most appropriate manner in the "Vorlesungen über technische Mechanik" by A. Föppl,[92] whose sixth volume (Leipzig 1910) also discusses in detail the theory of the top, and, in particular, the ship's gyroscope as a typical example of the theory presented by the author.

A short, summarizing, but comprehensive account of the theory of the top is given by A. G. Greenhill: Gyroscope and Gyrostat, Encyclopaedia Britannica, Suppl. 29, 1902.[93] With respect to the historical development of the theory of the top, we refer also to his lecture: The mathematical theory of the top, considered historically, proceedings of the 3rd International Mathematical Congress, Heidelberg 1904. See also the rich historical information in P. Stäckel's Encyklopädie-Artikel Bd. IV, 6; we thank the friendliness of Mr. Stäckel for a series of the following emendations.

To Chap. I

To p. 14.[94] The Euler angles first appeared in 1748, in Euler's Introductio in analysin infinitorum;[95] cf. P. Stäckel, Elementare Dynamik, Encykl. d. Math. Wiss. IV 6, Nr. 28a.

To p. 22, eqn. (10) and p. 57, eqn. (2). The transformation formulas (2) have been derived concisely in an elementary manner by A. Schoenflieſs[96] on the basis of the fact that every rotation in R_3 may be composed from two reversals (rotations by 180°), or also from two reflections. (Rend. del Circolo matematico di Palermo, t. 29, 1910.)

To p. 61. The vector product, as it commonly appears in vector analysis, is not identical with Graſsmann's "exterior product" (the manner of expression of the text, line 9 and below, is inexact in this respect), but rather with the Graſsmannian "complement of the exterior product." The vector product is an (axial) vector, while the exterior product, in contrast, is a "bivector" (that is, here, a surface element), but its complement (the normal erected upon the surface element with a given circulation sense) is an axial vector. The Graſsmannian theory proceeds through the enumeration of spatial quantities systematically; the physical vector calculus, opportunistically. The latter also abandons, in many cases, the difference in principle between polar vectors (vectors of the first kind) and axial vectors (vectors of the second kind), which is only permissible if only rotations of the coordinate systems, but not reflections or inversions, are used. Cf. Encyklopäedie der math. Wiss. IV, 14 (article by Abraham), Nr. 2, 3; R. Mehmke: Jahresbericht d. d. Math.ver. 13, 1904, p. 217; L. Prandtl: ibid. p. 436.[97] In the four-dimensional vector analysis of relativity theory used by Minkowski (see below), the difference between vectors of the first and second kinds emerges more sharply, in that the former have four components, while the latter have six components.[98]

To p. 64, eqn. (13). By means of eqn. (13), Hamilton had, a short time before the work of Cayley cited on p. 64, reduced rotations about a point in R_3 to quaternion multiplication. Proc. Irish Academy 11, Nov. 1844.[99]

This equation, by which we represent an arbitrary rotation-stretch of the x, y, z space about the starting point $x = y = z = 0$ in the most succinct form, is a special case of a more general representation that encompasses rotations of a four-dimensional space about a fixed point, or also, under another representation using homogeneous coordinates, the ∞^6 rotations and translations of ordinary three-dimensional space. For these transformation groups, the present generalization of eqn. (13) yields, at the same time, a very simple rational representation with 6 parameters (or 7, if rotation-stretches of R_4 are also included). The more general formula has, moreover, the advantage of greater symmetry and transparency.

Its simple basis is the often used theorem that the tensor of a quaternion product is equal to the products of the tensors of its constituents, where the tensor of a quaternion

$$Q = Ai + Bj + Ck + D$$

is understood to be the quantity

$$T = \sqrt{A^2 + B^2 + C^2 + D^2},$$

earlier designated also as the stretch. (This designation, introduced by Hamilton (cf. page 58), naturally has nothing to do with the now common concept of a tensor in mathematical physics.)

Let one quaternion be given as

$$v = xi + yj + zk + u,$$

and a second be defined through the quaternion product

$$\begin{gathered} (1) \qquad V = Xi + Yj + Zk + U = Q_1 v Q_2 \\ = (A_1 i + B_1 j + C_1 k + D_1)(xi + yj + zk + u)(A_2 i + B_2 j + C_2 k + D_2). \end{gathered}$$

Here the quantities X, Y, Z, U are obviously determined as linear homogeneous functions of the quantities x, y, z, u, and the relation

$$\begin{gathered} X^2 + Y^2 + Z^2 + U^2 = T_V^2 \\ = (A_1^2 + B_1^2 + C_1^2 + D_1^2)(x^2 + y^2 + z^2 + u^2)(A_2^2 + B_2^2 + C_2^2 + D_2^2) \\ = T_1^2 T_v^2 T_2^2 \end{gathered}$$

obtains.

We need now only require that

$$T_1^2 T_2^2 = 1 \tag{2}$$

in order that the linear transformation formulas between X and x represent an orthogonal transformation of the x, y, z, u space, which is indeed defined through the condition

$$X^2 + Y^2 + Z^2 + U^2 = x^2 + y^2 + z^2 + u^2. \tag{3}$$

The transformation is a rotation, since its determinant is $+1$, as one easily calculates afterward through detailed multiplication with Q_1 and Q_2. The transformation formulas at first contain the eight coefficients of the quaternions Q_1 and Q_2 as parameters, among which, however, the condition (2) obtains. Since, further, only the product of the quaternions Q_1 and Q_2 enters in eqn. (1), so that it remains unchanged if Q_1 is multiplied by a scalar factor λ and Q_2 is divided by this factor, one can choose this factor, for example, so that

$$\lambda^2 T_1^2 = \frac{1}{\lambda^2} T_2^2 = 1,$$

and then have the rotation of R_4 defined in the simplest manner by the six parameters of the two unit quaternions λQ_1, Q_2/λ. It is significant that this representation may be given immediately in rational form, as we will show below.

If ones lets, in contrast, the coefficients of Q_1 and Q_2 be arbitrary, then the quaternion product (1) contains seven parameters; in this case eqn. (1) leads to the "rotation-stretches of R_4." Formula (13) obviously results from this for the special specification

$$Q_2 = T_1 Q_1^{-1},$$

and yields, in addition, the transformation equations (2) of page 57 with the further relation $U = u$.

If we consider all the coefficients of the quaternions Q_1, Q_2, v, and V, and also λ, as real quantities, then real rotations in R_4 are represented by eqns. (1) and (2); we wish to designate them, in connection with a common manner of expression of projective measure geometry, as "elliptic" rotations. Next to these we place another group of collineations of R_4, the "hyperbolic" rotations, in which we set $u = \omega s$ and assume s as real, while $\omega = \sqrt{-1}$ (or, more generally, the square root of a negative quantity). Then the condition

$$X^2 + Y^2 + Z^2 + \omega^2 S^2 = x^2 + y^2 + z^2 + \omega^2 s^2 \tag{3$'$}$$

is fulfilled, which expresses that a real cone remains invariant for the transformation, which may even characterize the hyperbolic rotations.

To obtain real transformation formulas in this case, one must, as follows easily through a more detailed calculation, choose the two quaternions Q_1 and Q_2 as "conjugate," thus defining them as

$$\text{(4)}\qquad \begin{cases} Q_1 = (+Ai + Bj + Ck + D) + \omega(A'i + B'j + C'k + D'), \\ Q_2 = (-Ai - Bj - Ck + D) + \omega(A'i + B'j + C'k - D'). \end{cases}$$

The transformation formula would, more generally, also become real if Q_2 were to differ by a real factor from the expression given in (4); all possible transformations are thus divided into two types, according to whether this factor is positive or negative. Those of the former type form a group; for the latter, it is characteristic that $\partial S/\partial s$ is always negative; it is thus not possible that this second type can contain the identity substitution or infinitely small rotations. It will thus be advisable to restrict the definition of hyperbolic rotations to the first group, and therefore, by appropriate normalization, consider only formula (4). For the Lorentz transformation to be spoken of below, this implies the natural restriction that the time scales in all the considered coordinate systems should be conformably directed.

Without restriction of generality, a factor λ or $1/\lambda$ can also be chosen in (4) so that $T_1^2 = T_2^2$. Then if λQ_1 and Q_2/λ are also conjugate, λ must now be a complex quantity $a + \omega b$ of absolute value $\sqrt{a^2 - \omega^2 b^2} = 1$. If we imagine, for the sake of simplicity, that the quaternions Q_1 and Q_2 are already normalized, then the condition

$$\text{(4}'\text{)}\qquad AA' + BB' + CC' + DD' = 0$$

obtains.

It is now easy to see that we arrive, by means of quaternions, at a rational representation of rotations. Namely, to fulfill the condition (2) identically, we set, in place of (1), the transformation equation

$$\text{(1}'\text{)}\qquad (Xi + Yj + Zk + \omega S) = \frac{Q_1(xi + yj + zk + \omega s)Q_2}{T_1 T_2},$$

where now, however, because of (4′),

$$T_1 T_2 = T_1^2 = T_2^2 = A^2 + B^2 + C^2 + D^2 + \omega^2(A'^2 + B'^2 + C'^2 + D'^2),$$

thus becomes rational in the parameters A, A', ...

According to (4′), D', for example, is now, further, a rational homogeneous expression of the first degree in the seven remaining parameters; if one introduces it in the right side of (1′), then it becomes, since it is homogeneous of degree zero in the eight parameters of Q_1 and Q_2,

a mere function of the six ratios of A, A', B, B', C, C', D, and (1′) therefore leads, in fact, to a rational representation of hyperbolic rotations through six parameters.

These hyperbolic rotations now stand in close relation with the questions of modern physics to which the development of electrodynamics has led; this was the occasion for us to speak of them here. For the choice $\omega^2 = -c^2$, where c denotes the speed of light, we obtain a rational representation of the essential elements in the group that Poincaré has called the "Lorentz transformations," which is completed merely through the addition of an arbitrary displacement of the origin. These transformations play the same role, for electrodynamics and the associated questions, that the group, which consists of single rotations and displacements of the coordinate system and the same uniform translations, plays for classical mechanics, in that the fundamental electrodynamic equations are invariant if the Lorentz transformations are applied to the space coordinates x, y, z and the time coordinate s. This invariance makes possible a systematic development of the electrodynamics of moving bodies, as is claimed by the "relativity principle." Its physical meaning was recognized through the works of H. A. Lorentz, Versuch einer Theorie der elektrischen und magnetischen Erscheinungen in bewegten Köpern,[100] Leiden 1895 (2nd edition Leipzig 1906), and A. Einstein, Annalen der Physik 17, 1905, p. 891.[101] The group character of the Lorentz transformations was emphasized first by H. Poincaré, Rend. del circolo mat. di Palermo, 21, 1906, p. 129;[102] from the standpoint of the four-dimensional vector conception, finally, to which the representation above should yield a contribution, the development begins with H. Minkowski: Göttinger Nachrichten 1908, p. 53 and his lecture Raum und Zeit, Jahresbericht der deutschen Mathematikervereinigung 18, 1908.[103]

If one goes over in equations (4) to the limiting case $\omega^2 = 0$, and lets at the same time s and S become infinite of the order $1/\omega^2$, then one obtains a group of transformations for which $\omega^2 s = \omega^2 S$, and which can be designated as the "parabolic" rotations of R_4. Namely, if one interprets the transformation formulas in R_3, as whose coordinates one regards the quotients $x/\omega^2 s$, $y/\omega^2 s$, $z/\omega^2 s$, they then represent the group of ∞^6 linear orthogonal transformations of our usual (the parabolic) geometry, the rotations and translations. With the same mapping, the previously treated elliptic and hyperbolic cases lead to the linear orthogonal transformations of R_3, which correspond, respectively, to the

elliptic or hyperbolic metric specification; indeed, as is well known, one understands by these transformations the collineations for which an imaginary (eqn. 3) or a real (eqn. $3'$) surface of the second order remains invariant.

Finally, the limiting case in which ω^2 becomes infinite, and at the same time A', B', C', D' vanish to the order $1/\omega^2$, also leads to the "parabolic" rotations of R_4, while now x, y, z, s remain finite, and also $s = S$. One obtains a still further interpretation of this group, if one conceives of $s(= S)$ as the time coordinate, and, on the other hand, x, y, z and X, Y, Z as spatial coordinates. Namely, if one adds arbitrary displacements of the coordinate origin of the space and time measures, then there results the previously mentioned ∞^{10}-fold group under which Newtonian mechanics is invariant; it forms, in fact, the limiting case of the Lorentz group for infinitely large speed of light c. Contained as a subgroup, naturally, are the rotations in R_3, for which A', B', C', D' vanish identically, whereby our general equations (1) and (4) degenerate into the previous (13).

The equations (4), ($4'$), ($1'$) also encompass the previously treated "elliptic" rotations; namely, for positive values of ω^2, and therefore for real values of ω, for which the conjugate quaternions Q_1 and Q_2 become two arbitrary real quaternions. Without restriction of generality, one may therefore say that the characteristic subcases

$$\omega^2 = 1, \quad \omega^2 = -1, \quad \omega^2 = 0 \text{ or } \infty$$

directly represent the elliptic, hyperbolic, and parabolic rotations of R_4, or, respectively, the general linear orthogonal transformations of R_3 in the corresponding metric specification.

The application of quaternions to the representation of rotations was first noted by Cayley, under the restriction to the elliptical case (1). (On the homographic transformation of a surface of the second order onto itself, Philos. magazine VII, 1854, papers, t. II, p. 35, and Recherches ultérieures sur les déterminants gauches, Journal f. Math. t. 50, 1855, papers, t. II, p. 202.) The geometric interpretation was added by Klein (Zur nicht-euklidischen Geometrie, Math. Annalen 37, 1890). The general approach originates with Clifford (Preliminary sketch of biquaternions, Proceedings of the London Math. Society, Vol. IV, 1873, papers, London 1882, p. 181),[104] and is developed systematically by Study[105] (Von den Bewegungen und Umlegungen, II, Math. Annalen 39, 1891). For further literature, see the citations in Study, Encykl. d. Wiss. I, A, 4, in particular No. 35 of the French edition of the article (there I, 5).

To p. 68. We now prefer to interpret the concluding remark on “invariant thinking and invariant calculation” in accord with the significance that the vector calculus has in the meantime acquired (cf. also the note to page 142); not, in any case, to be understood in the sense that invariant vectorial calculation is something insignificant.

To Chap. III

To p. 138. It appears to us that the derivation of the Euler equations on pages 138 ff. may still not give, with the desirable transparency, the recognition that these equations are nothing other than the analytic expression of the impulse laws. For this reason the derivation follows here once again in a more succinct form, in which we use the manner of writing of vector calculus.

Because of the rotation R, an arbitrary fixed point in the top, whose coordinates x, y, z are denoted in the usual manner by the vector $\mathfrak{r}$ extended from the origin, has, according to the schema (3) on page 41, the velocity $U = V(R\mathfrak{r})$ in space, where the symbol V denotes, as on the page 138, the vector product defined on page 61.*) On the other hand, an arbitrary point $\mathfrak{r}'$ fixed in space moves with the velocity

$$U' = -V(R\mathfrak{r}') = V(\mathfrak{r}'R)$$

with respect to the body. Such a point is the impulse endpoint J in the force-free case. Its velocity relative to the body is thus

$$U' = \frac{dJ}{dt} = V(JR),$$

and thus follows eqn. (2) on page 140. If, on the other hand, the considered point is not fixed in space, but rather has the velocity W, it then has velocity $U' + W$ with respect to the body. For the impulse endpoint, this velocity W is equal, on the basis of the second impulse law, to the moment Δ of the forces; thus follows eqn. (3) on page 141:

$$\frac{dJ}{dt} = V(JR) + \Delta,$$

which, according to (3) and (3′), denotes precisely the system of the Euler equations in the vectorial conception.

The Euler equations are thus a pregnant example of a system of equations whose essential nature lies in its vector character. In general, vector theory is directly suitable to that form of the problem statement

*) Instead of $V(R\mathfrak{r})$, the notation $[R\mathfrak{r}]$ is now more common.

which we place at the summit of the treatment of rigid bodies: the problem of the movement of the impulse vector. That we abstained from the use of this type of calculation rests essentially on the fact that it has become adopted more and more only in the course of the last ten years, particularly proving, since then, to be an indispensable aid in theoretical electrodynamics. In its assessment, however, one should not overlook that for the actual treatment of the equations, and in particular for their integration, the passage to the coordinate equations is still generally required. In more recent works, the vector calculus is used in a consistent manner for the theory of the top; we cite Föppl, Lösung des Kreiselproblems mit Hilfe der Vektorenrechnung, Ztschr. f. Math. u. Phys., 48, 1903; Stübler, Der Impuls bei der Bewegung eines starren Körpers, ibid. 54, 1907.[106]

To p. 142. The simple interpretation of the Euler equations as the expression for the impulse theorems is found shortly before Hayward (1856) in Saint-Guilhem, Journ. de math. (1), 16, p. 347 (1851) and ibid. (1), 19, p. 346 (1854). For the history of the Euler equations, cf. Encykl. d. math. Wiss. IV, 6 (Art. Stäckel), No. 30.[107]

Felix Klein (seated in the center) and Arnold Sommerfeld (standing at the left) during a trip to England in 1899 (by courtesy of the Sommerfeld Project, Deutsches Museum, Munich).

Translators' Notes.

1. (Foreword, p. xvi) John Frederick William Herschel (1792–1871) was the only son of Friedrich Wilhelm (William) Herschel (1738–1822), the German-born discoverer of Uranus. The first supplementary note by Klein and Sommerfeld on page 197 identifies the source of the reference to the top as John Herschel's *Outlines of Astronomy* [Herschel 1851, pp. 190–191]. Herschel mentions the top in his discussion of the precessional motion of the Earth:

> The precession of the equinoxes thus conceived, consists, then, in a real but very slow motion of the pole of the heavens among the stars, in a small circle round the pole of the ecliptic. Now this cannot happen without producing corresponding changes in the apparent diurnal motion of the sphere, and the aspect which the heavens must present at very remote periods of history. The pole is nothing more than the vanishing point of the earth's axis. As this point, then, has such a motion as we have described, it necessarily follows that the earth's *axis* must have a conical motion, in virtue of which it points successively to every part of the small circle in question. We may form the best idea of such a motion by noticing a child's peg-top, when it spins not upright, or that amusing toy the te-to-tum, which, when delicately executed, and nicely balanced, becomes an elegant philosophical instrument, and exhibits, in the most beautiful manner, the whole phenomenon.

A te-to-tum (or, now more commonly, teetotum) is a solid top, twirled by the fingers, that spins on a sharp point. The body of the top is often inscribed with numbers or letters for use in games. Examples from around the world are illustrated in Douglas W. Gould's interesting book *The Top* [Gould 1973, p. 44].

The Foreword, which faces more to the rear than the fore, was written when the fourth and final volume of the *Theorie des Kreisels* was published in 1910. The following Advertisement (*Anzeige*) appeared shortly before Vol. I was published in 1897.

2. (Advertisement, p. xviii) The German Association for the Advancement of Mathematical and Scientific Education (*Deutsche Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts*) was founded in 1891, and exists today as one of the largest teachers' associations in Germany. Klein's 1895/96 lecture on the top in Göttingen was given to members of this association.

3. (page 1) Constant Marie Rozé (1840–1911) was professor of mathematics and mechanics in *l'École municipale de Physique et Chemie industrielles* in Paris from 1882–1910. He worked primarily in the field of precision chronometry. No mention of Rozé outside of the present work by Klein and Sommerfeld has been found by the translators in the literature on tops and gyroscopes.

4. (page 2) An apparatus for spinning a bell-shaped top is shown in Fig. 144. The figure is reproduced from the beautifully illustrated *Physikalische Demonstrationen* by Adolph Ferdinand Weinhold (1841–1917), a teacher in the *königliche höhe Gewerbschule* in Chemnitz. After unwinding the silken cord by turning the crank E, the rotating top and its pedestal are released by folding piece C to the right on the hinges cc, and then folding piece A to the rear on its hinged connection to D. Weinhold states that the top will spin on its agate seat for one hour in air, and for two hours in an evacuated chamber. The top and spinning apparatus appear in the early twentieth-century catalogues of Max Kohl AG, a manufacturer of scientific instruments in Chemnitz from 1876 to 1946.

5. (page 2) Johann Gottlob Friedrich von Bohnenberger (1765–1831) was professor of mathematics and astronomy in the University of Tübingen. A drawing of the original Bohnenberger machine is shown in Fig. 145. Bohnenberger writes that the machine "was made very precisely and elegantly by Herr Universitäts-Mechanikus Buzengeiger for the price of 18 Gulden," and gives a detailed description of its use in illustrating the rotational and precessional motion of the Earth.

The dual-ring apparatus used to support the rotating sphere in Fig. 145 is sometimes called *Cardan's suspension* [Synge 1949, p. 418]. In an article on dynamics in the *Encyklopädie der mathematischen Wissenschaften*, Paul Gustav Stäckel (1862–1919; professor of mathematics in Heidelberg) cites a sixteenth-century work of Girolamo Cardano (1501–1576; also Hieronymus Cardanus or Jerome Cardan) in which Cardano describes the dual-ring suspension as an "old invention" [Stäckel 1908, p. 559].

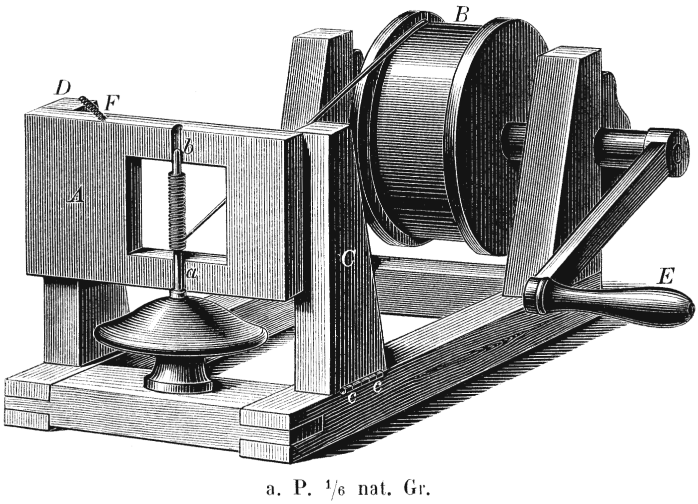


Fig. 144. Mechanism for spinning a bell-shaped top [Weinhold 1881, p. 58].

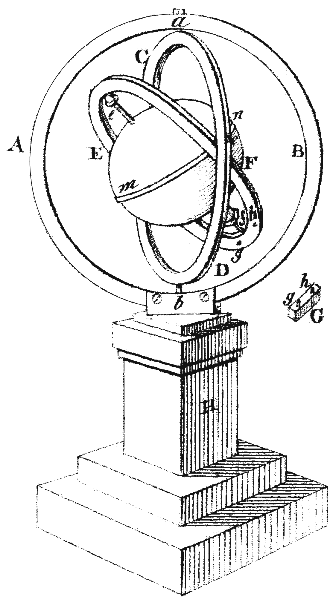


Fig. 145. Bohnenberger machine for illustrating the rotational and precessional motion of the Earth [Bohnenberger 1818, Taf. I].

6. (page 2) Jean Bernard Léon Foucault (1819–1868) used his gyroscope (the word is due to Foucault) to give the first nonastronomical experimental demonstration of the rotation of the Earth. Klein and Sommerfeld discuss Foucault's experiments in Vol. III, Chapter VIII, § 9.

7. (page 4) The famous statement of Joseph-Louis Lagrange (1736–1813) appears in the concluding paragraph of the *Avertissement* of the first edition of the *Méchanique Analitique* [Lagrange 1788, p. vj]:

> No figures will be found in this work. The methods that I present require neither constructions nor geometric or mechanical arguments, but only algebraic operations subjected to a regular and uniform march. Those who love analysis will see with pleasure mechanics becoming a new branch, and will be grateful to me for having thus extended its domain.

8. (page 5) William Thomson (1824–1907; 1st Baron Kelvin of Largs, 1892) and Peter Guthrie Tait (1831–1901) use the term Natural Philosophy "to denote the investigation of laws in the material world, and the deduction of results not directly observed. Observation, classification, and description of phenomena necessarily precede Natural Philosophy in every department of natural science. The earlier stage is, in some branches, commonly called Natural History; and it might with equal propriety be so called in all others" [Thompson 1879a, p. v]. In the second part of the second (1883) edition of the *Treatise on Natural Philosophy*, Thomson and Tait write that "it was intended that the various branches of mathematical and experimental physics should be successively treated. The intention of proceeding with the other volumes is now definitely abandoned" [Thompson 1883, p. v]. The 1912 edition of the *Treatise* was reissued by Dover Publications in 1962 under the title *Principles of Mechanics and Dynamics*.

9. (page 5) Edward John Routh (1831–1907) placed first in the 1854 Mathematical Tripos examination in Cambridge University; James Clerk Maxwell (1831–1879) finished second. Klein and Sommerfeld visited Routh during a trip to England in 1899.

10. (page 5) Jean Marie Constant Duhamel (1797–1872), *Cours de mécanique de l'École polytechnique* [Duhamel 1845]; Théodore Despeyrous (1815–1883) and Gaston Darboux (1842–1917), *Cours de mécanique; avec des notes par m. G. Darboux* [Despeyrous 1884];

Paul Émile Appell (1855–1930), *Traité de mécanique rationelle* [Appell 1893].

11. (page 5) Woldemar Voigt (1850–1919) became professor of physics in Göttingen in 1883. Sommerfeld was once offered a postdoctoral assistantship with Voigt, but declined the offer to become the assistant of Felix Klein [Eckert 2003, p. 169].

12. (page 5) Emil Arnold Budde (1842–1921) was a student of Julius Plücker (1801–1868), who held professorial chairs in both physics and mathematics in the University of Bonn. Budde received the Ph.D. degree in physics in 1864; Felix Klein was an assistant of Plücker in mathematics from 1866 to 1868. In 1893 Budde became the director of the Charlottenburg factory of Siemens & Halske AG, the forerunner of the modern German electronics, power, and communications company.

13. (page 5) The *Mechanik* of Gustav Robert Kirchhoff (1824–1887) is Vol. I of the four-volume *Vorlesungen über mathematische Physik* [Kirchhoff 1876].

14. (page 5) The life of Louis Poinsot (1777–1859) is summarized in an obituary notice that appeared in the Proceedings of the Royal Society of London, Vol. XI, 1860–1862:

> M. Louis Poinsot was born in Paris on the 3rd of January, 1777; quitted the École Polytechnique in the capacity of Ingénieur des Ponts et Chaussées in 1796; was appointed Professor at the Lycée Bonaparte; Professor and Examiner of the pupils who had completed the course of instruction at the École Polytechnique, and member of the Council for improving that institution. He was elected Member of the Institute in 1813, in the place rendered vacant by the death of Lagrange, by whom his future eminence had been predicted. He became Grand Officer of the Legion of Honour, Peer of France and Senator. He died, unmarried, on the 5th of December, 1859, at the age of nearly 83 years. The greater part of his life passed happily; while still young his talents were appreciated by the most eminent judges; in his advanced years he received the highest rewards from his countrymen. With his simple habits, the emoluments of the numerous and honourable appointments he held, and the profits of nine large editions of his 'Statics,' in addition

> to a moderate family inheritance, enabled him to leave at his death a fortune of upwards of 1,200,000 francs.

Poinsot's *Théorie nouvelle de la rotation des corps* first appeared in 1834 [Poinsot 1834], and a second edition was published in 1851 [Poinsot 1851]. The 1834 edition of the work consists of fifty-six pages of text, with no equations and no figures. After describing a purely geometric solution of the general problem of rigid body motion, Poinsot gives his opinion of analytic mechanics in a section entitled *Réflexion général*:

> We are thus led by reasoning alone to a clear conception that the geometers were not able to extract from the formulas of analysis. This is a new example that shows the advantage of the simple and natural method of considering things in themselves, without losing them from view in the course of the reasoning. For if one is content, as one usually is, to translate problems into equations, and then rely on transformations by calculations to bring to light the solution that one has in mind, one will most often find that this solution is yet more concealed in analytic symbols than it was in the very nature of the proposed question. The art of discovery resides not in calculations; but in that attentive consideration of things where the mind seeks, above all, to form an idea, trying, by analysis in its proper sense, to decompose things into other simpler ones, in order to review them later as if they were formed by the reunion of these simple things of which the mind has a full knowledge. It is not that things are constituted in this manner, but it is our only way to see them, to form an idea, and thereby know them. Thus our true method is only this happy mix of analysis and synthesis, where calculation is employed only as an instrument. A precious and necessary instrument, without doubt, because it ensures and facilitates our progress; but which does not have in itself any intrinsic virtue; which does not direct the mind, but which the mind must direct, as all other instruments.
>
> What may have created an illusion in some minds of this kind of power which they attribute to the formulas of analysis, is that one recoups, with some facility, truths already known, and which one has, so to speak, introduced

oneself; and it then seems that analysis gives us that which it only returns to us in another language. When a theorem is known, one has only to express it by equations; if the theorem is true, each of those cannot avoid being exact, as well as the transformations that one can deduce from them: and if one arrives in this way to some evident or otherwise well-established formula, one has only to take this expression as a point of departure, to trace back, and the calculation alone appears to have brought us, as if by itself, to the theorem in question. But it is here that the reader is deceived. Thus, to take an example in the very question that is the object of this memoir, it is quite clear that nothing would be easier today than to recover our ideas in the analytic expressions of Euler or Lagrange, and even to derive them with an air of facility that would make one believe that these formulas must produce them spontaneously. However, since these ideas have until now escaped so many Geometers who have transformed these formulas in so many ways, it must be admitted that this analysis did not yield them at all, since to see them one had to wait until another arrived by a completely different path.

We would have many other reflections to make, and greater examples to produce, if we wished to show, on the one hand, all that the mind owes for enlightenment to this natural method, as I have defined it above, and which constitutes our true analysis, and, on the other hand, the few new truths that one has been able to extract from these analytic formulas in which one thinks to enclose a question, and sometimes even an entire science. Without doubt, science is contained there, as it would be in any principle enunciated in general terms; but the difficulty remains to make it emerge; and has not this difficulty become greater? And, for instance, must one not know both Mechanics and Analysis to extract from the single general formula of virtual velocities not, say, a new theorem (of which I see few examples), but only the particular propositions that are already the most well known? Is not the translation here more difficult than the text itself, which, I mean to say, is the immediate consideration of the things that one would study? The illustrious author who would transform

Mechanics into a question of calculations has, without a doubt, attained his objective with all the clarity and elegance that one could expect. But if true analysis shines somewhere in the *Mécanique analytique*, I would dare to say that it is much less in the calculations that the author organizes with such order and symmetry, than in the *rapprochement* of methods, and in the admirable prefaces that he has placed at the heads of the different books of his work, where he examines and discusses the fundamental principles of the science, and which give an instructive history of the progress of the human mind in this delicate sequence of subtle ideas and ingenious solutions that have little by little formed Mechanics. It is especially there that this beautiful work will be able to serve the further progress of the mind, by showing it the path on which it must continue to march. Because, once again, let us guard against believing that a science is complete when one has reduced it to analytic formulas. Nothing excuses us from studying things in themselves, and being well aware of the concepts that are the object of our speculations. Let us not forget that the results of our calculation must almost always be verified from another side, by some simple reasoning, or by experience. That if calculation only can sometimes offer us a new truth, one must not believe at this very point that the mind does not have anything more to do: but, on the contrary, one must consider that this truth is independent of the methods or artifices that may have led us to it, and that there certainly exists a simple demonstration that could make it evident: this must be the great objective and the final result of mathematical science.

Forgive these reflections, which I make, I dare say, only in the interest of science. I know the intrinsic and distinctive character of algebraic analysis, and I could even say with precision how this art has been able to perfect the ordinary logic of discourse: I know what all good minds owe to calculation; but I try to enlighten those who are deceived in the nature of this instrument, and at the same time try to prevent the abuse that others have committed by taking advantage of this very illusion. Because, as soon as an ingenious author has managed to reach a new truth, should

> not one fear that the most sterile calculator will hurry to find it in his formulas, to discover it a second time, and in his own manner, which he says to be the correct and true; in such a way that one believes oneself indebted more to his analysis, and the author himself, sometimes not well versed or even foreign to this language and these symbols with which his ideas are stolen, barely dares to claim what belongs to him, and withdraws almost confused, as if he had poorly invented what he had discovered so well? A singular artifice, which I need not characterize more, but which is well to point out as one of the most harmful to scientific progress, because it is without exception one of the most likely to discourage inventors!
>
> But I will extend these reflections no further: and if the little that has been said is sensible enough by the preceding examples, one will see it confirmed again by those that follow.

By the time that the second edition of the *Théorie nouvelle de la rotation des corps* was published in 1851, Poinsot appears to have moderated his views; the one hundred and sixty-seven pages of the second edition are as full of formulas as any textbook of analytic mechanics. The now incongruous *Réflexion général*, however, is retained in its entirety.

15. (page 7) Thomson and Tait attribute their use of the word *kinematics* to an unreferenced suggestion of André Marie Ampère (1775–1836). In the *Elements of Natural Philosophy*, they introduce the word *kinetics* as follows [Thomson 1879b, p. 1]:

> 1. The science which investigates the action of Force is called, by the most logical writers, DYNAMICS. It is commonly, but erroneously, called MECHANICS; a term employed by Newton in its true sense, the Science of Machines, and the art of making them.
>
> 2. Force is recognized as acting in two ways:
>
> 1° so as to compel rest or to prevent change of motion, and
>
> 2° so as to produce or to change motion.
>
> Dynamics, therefore, is divided into two parts, which are conveniently called STATICS and KINETICS.

16. (page 8) The construction for the displacement and rotation that bring the noncollinear points O, P, Q from their initial position $O_1P_1Q_1$ to their final position $O_2P_2Q_2$ is illustrated in Fig. 146. The rotation axis R passes through O_2, and is in the direction of the unit vector

$$\mathbf{u}_R = \frac{\overrightarrow{Q_1'Q_2} \times \overrightarrow{P_1'P_2}}{\left|\overrightarrow{Q_1'Q_2} \times \overrightarrow{P_1'P_2}\right|}.$$

The rotation angle ϑ can be obtained by first computing the vectors

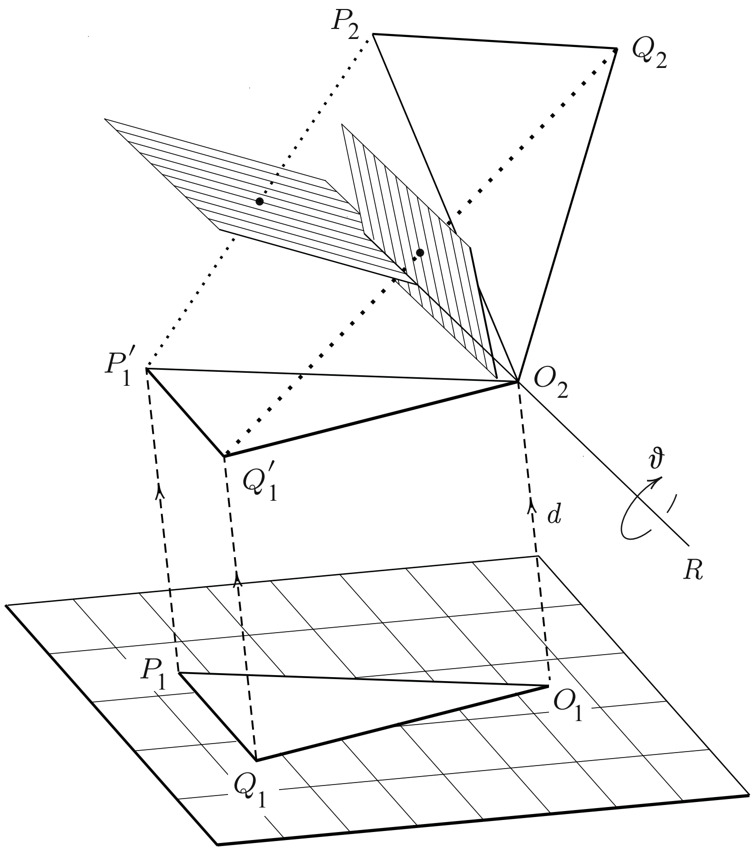


Fig. 146. Construction for the determination of the rotation axis and the rotation angle corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$.

$$\mathbf{n}_1 = \overrightarrow{O_2Q_2} - (\overrightarrow{O_2Q_2} \cdot \mathbf{u}_R)\mathbf{u}_R, \quad \mathbf{n}_2 = \overrightarrow{O_2Q_1'} - (\overrightarrow{O_2Q_1'} \cdot \mathbf{u}_R)\mathbf{u}_R,$$
$$\mathbf{n}_3 = \overrightarrow{O_2P_2} - (\overrightarrow{O_2P_2} \cdot \mathbf{u}_\mathbf{R})\mathbf{u}_R, \quad \mathbf{n}_4 = \overrightarrow{O_2P_1'} - (\overrightarrow{O_2P_1'} \cdot \mathbf{u}_R)\mathbf{u}_R$$

in the plane perpendicular to the rotation axis R. Then

$$\cos\vartheta = \frac{\mathbf{n}_1 \cdot \mathbf{n}_2}{|\mathbf{n}_1||\mathbf{n}_2|} = \frac{\mathbf{n}_3 \cdot \mathbf{n}_4}{|\mathbf{n}_3||\mathbf{n}_4|}.$$

The direction of the rotation axis and the rotation angle ϑ are independent of the choice of the reference point O.

The translation and rotation corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$ are illustrated in Fig. 147.

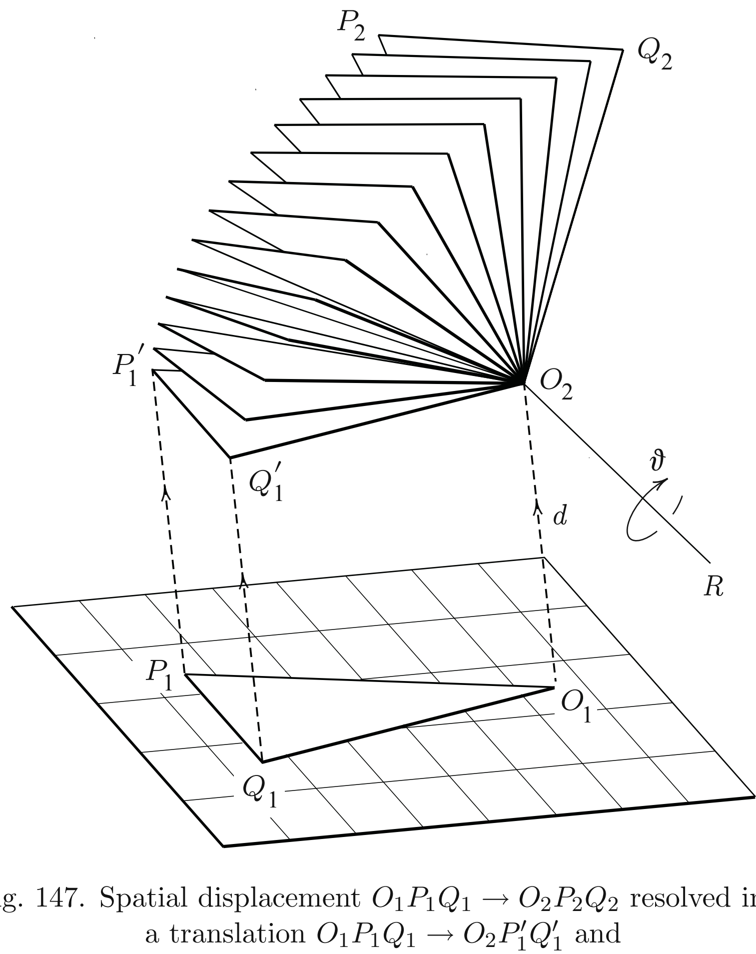


Fig. 147. Spatial displacement $O_1P_1Q_1 \to O_2P_2Q_2$ resolved into a translation $O_1P_1Q_1 \to O_2P_1'Q_1'$ and a rotation $O_2P_1'Q_1' \to O_2P_2Q_2$.

17. (page 8) A rotation-pair is a rotation about a body axis followed by an equal but oppositely directed rotation about a parallel body axis [Schell 1879, p. 168]. If the distance between the axes is d and the angle of rotation is ϑ, the magnitude of the equivalent translation is $2d \sin \vartheta/2$. The translation is in the direction of the chord of the circular arc that an arbitrary point on one of the axes makes during the sequence of the two rotations. Interchanging the order of the rotations reverses the direction, but does not change the magnitude of the translation.

18. (page 9) The corresponding theorem for the statics of a rigid system is discussed by Klein and Sommerfeld in Chapter II, §2.

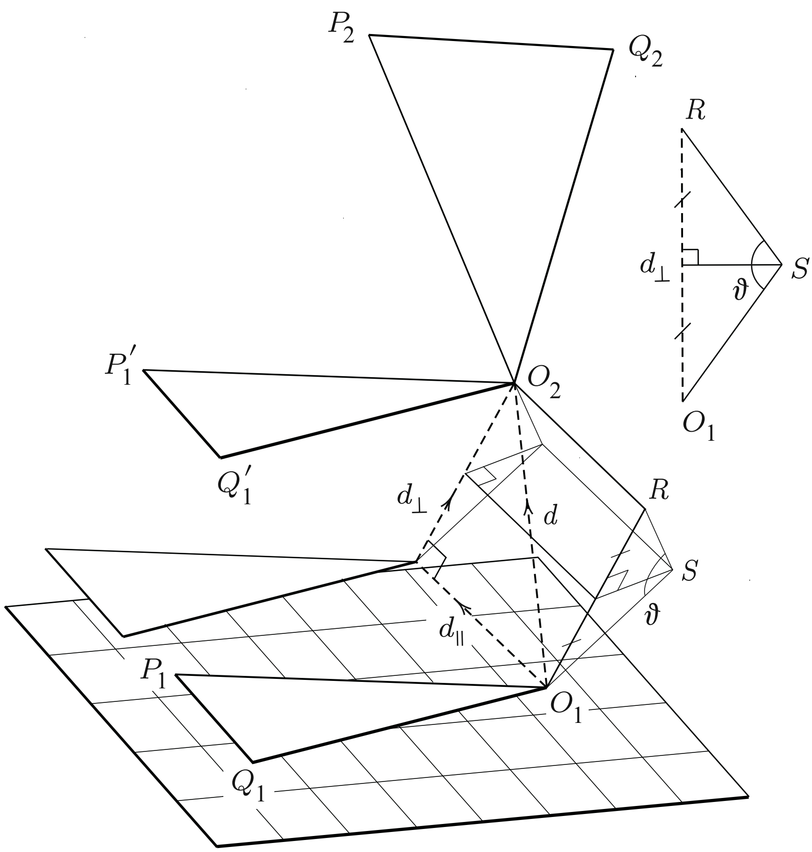


Fig. 148. Construction for the determination of the screw parameters corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$.

19. (page 9) The construction for the determination of the parameters of the motion-screw corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$ is illustrated in Fig. 148. To find the axis of the screw, the displacement d in Fig. 146 is first decomposed into a component $d_{\|}$ parallel to the rotation axis and a component $d_{\perp}$ perpendicular to the rotation axis. The screw axis S then lies in the normal plane that bisects $d_{\perp}$. The distance from $d_{\perp}$ to S is $(d_{\perp}/2)\tan\vartheta/2$, where ϑ is the rotation angle determined in Fig. 146.

The screw motion corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$ is illustrated in Fig. 149.

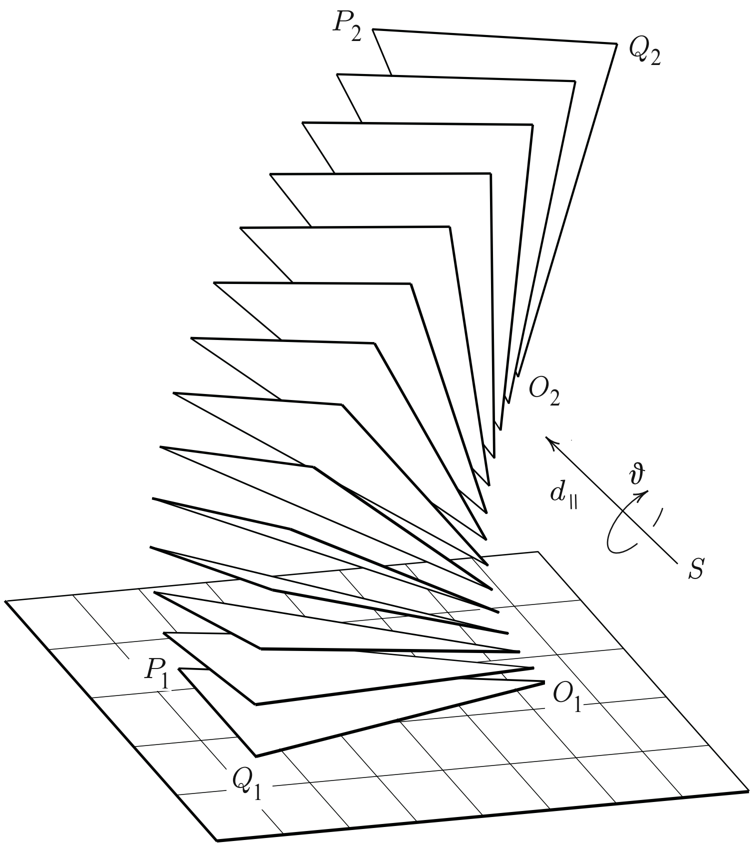


Fig. 149. Screw motion corresponding to the displacement $O_1P_1Q_1 \to O_2P_2Q_2$.

20. (page 10) A three-sided corner is formed by three planes that meet at a point O; a familiar example is the corner of a cube. Such a corner intersects a sphere with center at the vertex O in a spherical triangle ABC, as shown in Fig. 150. The edge angles of the corner are the angles α, β, and γ of the spherical triangle. Rotations of space about the edges of the corner may be represented by rotations of the sphere about the fixed axes OA, OB, and OC.

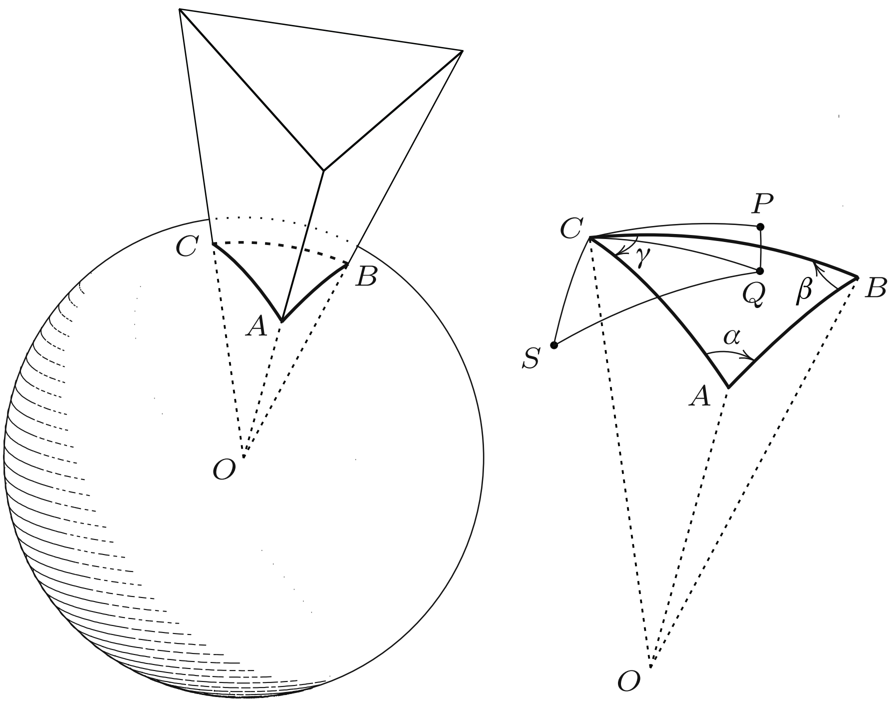


Fig. 150. Three-sided corner and its intersection with a unit sphere.

An elegant proof of the three-rotation theorem is given by John H. Conway and Derek A. Smith [Conway 2003, p. 26]. Suppose that the sphere in Fig. 150 is first reflected about the great circle containing the arc BC and then reflected about the great circle containing the arc AC, so that a point P on the sphere is reflected first to point Q and then to point S. The point C and its antipodal point are the only points on the sphere that are unchanged in the sequence of the two reflections, and the two reflections must therefore be equivalent to a rotation about the axis OC. The angle of rotation is equal, from the

figure, to 2γ. If $\mathbf{R}(OC, 2\gamma)$ represents the rotation by 2γ about OC, $\boldsymbol{\text{Я}}(BC)$ represents the reflection about arc BC, and $\boldsymbol{\text{Я}}(AC)$ represents the reflection about arc AC, then

$$\mathbf{R}(OC, 2\gamma) = \boldsymbol{\text{Я}}(AC)\boldsymbol{\text{Я}}(BC).$$

Similarly,

$$\mathbf{R}(OB, 2\beta) = \boldsymbol{\text{Я}}(BC)\boldsymbol{\text{Я}}(AB) \quad \text{and} \quad \mathbf{R}(OA, 2\alpha) = \boldsymbol{\text{Я}}(AB)\boldsymbol{\text{Я}}(AC).$$

The theorem now follows, for

$$\begin{aligned} &\mathbf{R}(OC; 2\gamma)\mathbf{R}(OB; 2\beta)\mathbf{R}(OA; 2\alpha) \\ &\qquad = \boldsymbol{\text{Я}}(AC)\boldsymbol{\text{Я}}(BC)\boldsymbol{\text{Я}}(BC)\boldsymbol{\text{Я}}(AB)\boldsymbol{\text{Я}}(AB)\boldsymbol{\text{Я}}(AC) \\ (1) \qquad &\qquad = \mathbf{I}, \end{aligned}$$

where, since the product of a reflection with itself leaves space unchanged, $\mathbf{I}$ is the identity operator.

In his *Lectures on Quaternions*, Hamilton states a more general form of the theorem [Hamilton 1853, p. 334]:

> If a *body* B be made to *revolve* through *any number of successive and finite rotations*, represented as to their axes and amplitudes by the DOUBLES OF THE ANGLES, A_1, A_2, ..., A_n, *of any spherical polygon*, this body B will be BROUGHT BACK, hereby, *to its own original position.*

Hamilton also makes a curious statement regarding his right to claim primacy in the discovery of the theorem:

> You will find, by the printed Proceedings of the Royal Irish Academy, that I stated this theorem (with only a slight difference in its wording), at a general meeting of that Academy, in November, 1844, as a consequence of those principles respecting Quaternions, which had been communicated to the Academy by me, about a year before. The theorem, at that time, appeared to me to be new; nor am I able, at this moment, to specify any work in which it may have been anticipated: although it seems to me *likely* enough that some such anticipation may exist. Be that as it may, the theorem was certainly *suggested to me* by the quaternions; nor can I easily believe that *any other* mathematical method shall be found to furnish any SIMPLER form of EXPRESSION for the same general geometrical result.

21. (page 10) The geometric construction for the determination of the rotation axis and the rotation angle corresponding to the composition of two finite rotations is illustrated in Fig. 151. Suppose that a rotation about axis OA with rotation angle α is followed by a rotation about axis OB with rotation angle β. After the construction of the spherical triangle in Fig. 151(a), it follows from equation (1) on the previous page that

$$\mathbf{R}(OC, 2\psi)\mathbf{R}(OB, \beta)\mathbf{R}(OA, \alpha) = \mathbf{I},$$

so that

$$\begin{aligned}\mathbf{R}(OB, \beta)\mathbf{R}(OA, \alpha) &= \mathbf{R}^{-1}(OC, 2\psi) = \mathbf{R}(OC, -2\psi)\\ &= \mathbf{R}(OC, -2(\pi - \gamma/2)) = \mathbf{R}(OC, -2\pi + \gamma) = \mathbf{R}(OC, \gamma).\end{aligned}$$

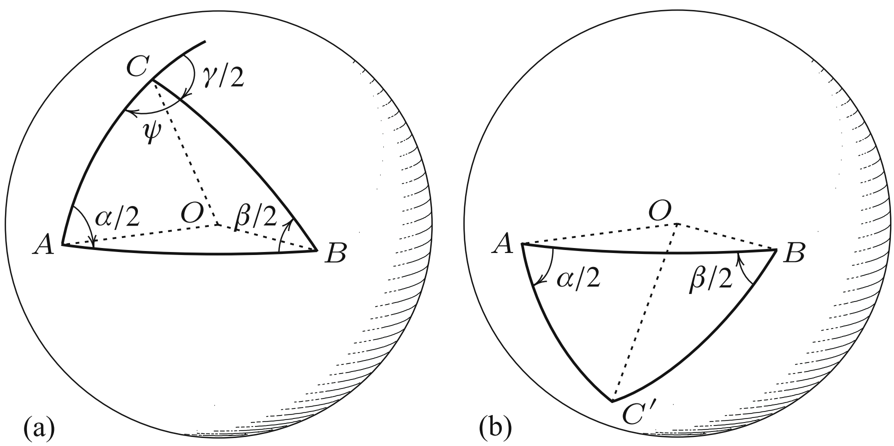


Fig. 151. Geometric construction for the rotation axis and rotation angle corresponding to the composition of two finite rotations.

According to Simon L. Altmann [Altmann 1989, p. 302], this construction for the resultant of two finite rotations was first discovered by Benjamin Olinde Rodrigues (1794–1851). Rodrigues states his result in the following remarkable way [Rodrigues 1840, pp. 390–391]:

> The resultant axis must be placed in such a way that, having accomplished the two rotations about the two given convergent axes, it returns to its first position. If, therefore, one passes through each of the given axes a plane that makes, with respect to the plane of the two given axes, an angle equal to the half-rotation corresponding to that axis,

> the intersection of these two planes will be the desired resultant axis, arriving by virtue of the first rotation to its *symmetric* point with respect to the plane of the two given axes, and returning in the second rotation to its original position.
>
> One sees, at the same time, that the angle between these two planes will measure the half-rotation of the composed rotation, because the first axis, immobile during the first rotation, will not be displaced except by the second rotation, and will describe around the resultant axis, determined as has just been described, an angle twice that between the two planes. Let us observe here that the half-rotation about each axis can be measured indifferently by the interior or the exterior angle of the two planes that pass through this axis, the sense of the rotation only being changed, depending on whether one adopts one or the other measure, since any rotation accomplished in a definite sense about an axis is equivalent to a contrary rotation of amplitude complementary to the first with respect to 400°.

The division of the circle into 400 degrees (and thus the right angle into 100 degrees) was one of the metric reforms adopted during the French Revolution. According to Thomson and Tait, "The decimal division of the right angle, decreed by the French Republic when it successfully introduced other more sweeping changes, utterly and deservedly failed" [Thomson 1879a, p. 459]. Not quite utterly; the gradian (German *Neugrad* or *Gon*), defined as the hundredth part of the right angle, is still used occasionally today, and appears on many calculators.

22. (page 11) In their very fine book *Gyrodynamics*, Ronald N. Arnold and Leonard Maunder show how the geometric construction for the resultant of two finite rotations reduces to the parallelogram law as the two rotations become infinitesimal small [Arnold 1961, pp. 42–44].

23. (page 11) If the rotation with angle β about the axis OB is followed by the rotation with angle α about the axis OA, the resultant rotation has the axis OC' shown in Fig. 151(b). The point C' in Fig. 151(b) is the reflection of the point C in Fig. 151(a) with respect to the plane AOB. The angle γ of the resultant rotation is independent of the order of the rotations about OA and OB.

24. (page 12) *Vielkant.*

25. (page 13) *schroten.*

26. (page 14) In the original (1834) edition of his *Théorie nouvelle de la rotation des corps*, Poinsot uses the French names *serpoloïde* and *poloïde* for the two curves. In the 1851 edition of the *Théorie nouvelle*, the names are changed to *herpolhodie* and *polhodie*, and a Greek etymology is added.

27. (page 15) Maxwell's two-page paper is entitled "On an Instrument to illustrate Poinsôt's Theory of Rotation" [Maxwell 1965, pp. 246–247]. A later paper by Maxwell [Maxwell 1965, pp. 248–264] contains a drawing, reproduced in Fig. 152, of a top with the colored disk that is used to observe the instantaneous axis of rotation. Maxwell writes that

> The best arrangement, for general observation, is to have the disc of card divided into four quadrants, coloured with vermilion, chrome yellow, emerald green, and ultramarine. These are bright colours, and, if the vermilion is good, they combine into a grayish tint when the revolution is about the axle, and burst into brilliant colours when the axis is disturbed. It is useful to have some concentric circles, drawn with ink, over the colours, and about 12 radii drawn in strong pencil lines. It is easy to distinguish the ink from the pencil lines, as they cross the invariable axis, by their want of lustre. In this way, the path of the invariable axis may be identified with great accuracy, and compared with theory.

The screws on the body of the top in Fig. 152 are used to adjust the position of the center of gravity of the top, the magnitudes of the principal moments of inertia, and the directions of the principal axes. A detailed analysis of both the screw adjustments and the motion of the Maxwell top is given by Max Winkelmann (1879–1945) in his 1904 Göttingen dissertation [Winkelmann 1904].

28. (page 15) *gleichstimmig.*

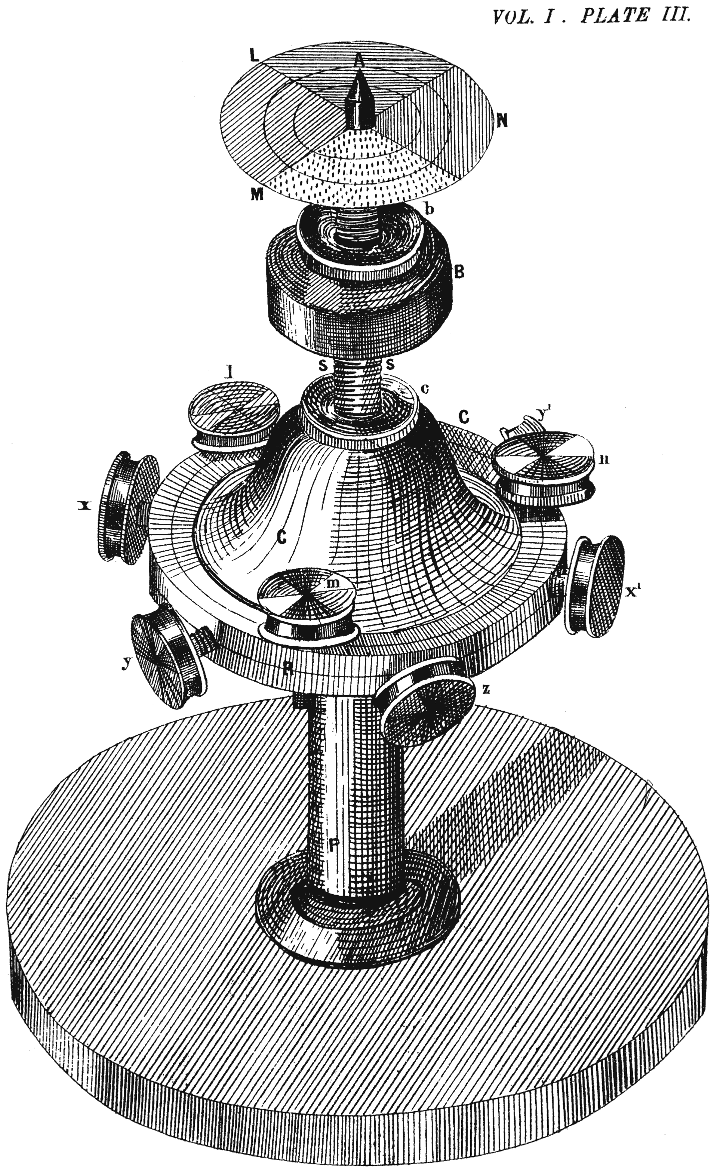


FIG. 6.

Fig. 152. Maxwell's top, with pasteboard disk LMN for the visualization of the path of the instantaneous axis of rotation [Maxwell 1965, p. 264].

29. (page 17) If the direction cosines are grouped into the matrix

$$\mathbf{C} = \begin{bmatrix} a & a' & a'' \\ b & b' & b'' \\ c & c' & c'' \end{bmatrix},$$

then the constraints among the direction cosines are contained in the equations

$$\det \mathbf{C} = 1, \quad \mathbf{C}\mathbf{C}^T = \mathbf{I},$$

where $\mathbf{I}$ is the 3×3 identity matrix. These equations imply that each element of $\mathbf{C}$ is equal to its cofactor.

30. (page 18) According to the list of corrections to Vol. I that appeared when Vol. IV was published in 1910, Fig. 3 on p. 18 "is a regrettable failure in perspective." A clearer illustration of the Euler angles is given in Fig. 153.

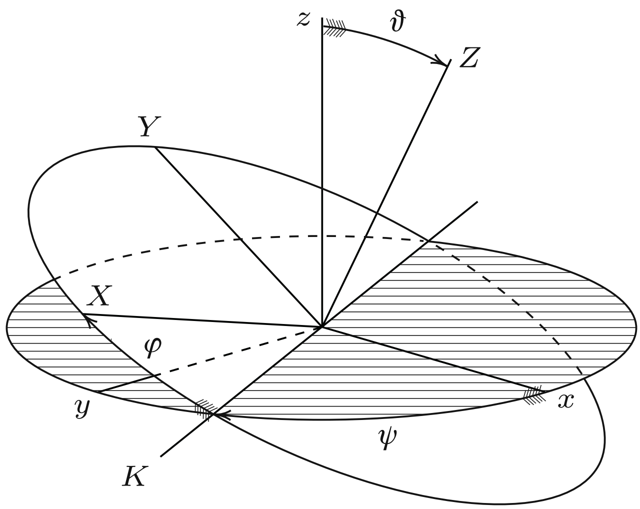


Fig. 153. Fixed spatial frame xyz, top frame XYZ, line of nodes K, and Euler angles φ, ψ, ϑ [Winkelmann 1904, p. 22].

31. (page 24) The complex quantities λ_1 and λ_2 in eqn. (3) on page 24 are the two components of what is now called a spinor. The general theory of spinors for the representation of the group of orthogonal transformations of n-dimensional space (and other semisimple groups) was developed by Elie Cartan (1869–1951) [Cartan 1913; Brauer 1935]. Spinors have found wide application in quantum mechanics; see, for

example, *The Story of Spin*, by the Nobel laureate Sin-itiro Tomonaga [Tomonaga 1997].

32. (page 25) The quantities α, β, γ, δ are now called the Cayley–Klein parameters; the matrix

$$\begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$$

is an element of the group SU(2). Klein first used the homographic transformation (5) on page 25 to study spatial symmetry groups, and, in particular, the symmetry group of the icosahedron [Klein 1875; Klein 1877]. Arthur Cayley (1821–1895) derived the axis and angle of rotation, the Rodrigues parameters, and the quaternion quantities that correspond to the coefficients of a given homography [Cayley 1879].

33. (page 28) The tangent plane to the minimal cone is

$$\hat{\xi}\eta + \hat{\eta}\xi = 2\hat{\zeta}\zeta,$$

where (ξ, η, ζ) is a point in the tangent plane and $(\hat{\xi}, \hat{\eta}, \hat{\zeta})$ is a point on the cone. Introducing the parameter

$$\lambda = \frac{\hat{\xi}}{\hat{\zeta}} = \frac{\hat{\zeta}}{\hat{\eta}},$$

which distinguishes the generators of the cone, the equation of the tangent plane can be written as

$$\xi + \lambda^2\eta = 2\lambda\zeta.$$

The two parameters λ that corresponds to the point (ξ, η, ζ) are therefore

$$\begin{aligned} \lambda &= \frac{\zeta \pm \sqrt{\zeta^2 - \xi\eta}}{\eta} \\ &= \frac{\zeta \pm \sqrt{\zeta^2 - (\zeta^2 - r^2)}}{\eta} = \frac{\zeta \pm r}{\eta}, \end{aligned}$$

in agreement with equations (6) and (6′) on pages 25 and 26, respectively.

34. (page 36) This paragraph gives a rare and extraordinary insight into the creative processes of three first-rate scientists. Klein discovered spinors in his study of the icosahedron. Cayley found the relation

between the groups SU(2) and SO(3). Klein and Sommerfeld here puzzle over the fact that SU(2) is a double cover of the rotation group SO(3). The intermediate states of Klein and Sommerfeld are hidden in the spinor space. The spinor changes its sign as it rotates by 2π, but returns to its initial value after a rotation of 4π; hence the two possible conventions for the signum of the Cayley–Klein parameters.

35. (page 54) The polhode and herpolhode cones and curves for uniformly accelerated precession are illustrated in Fig. 154.

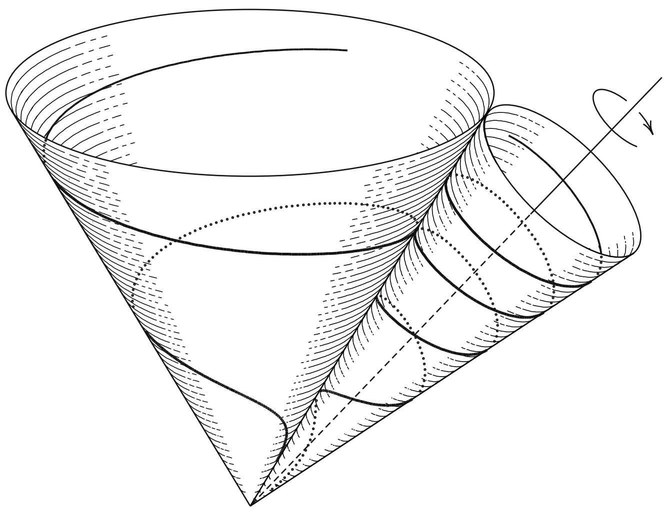


Fig. 154. Polhode and herpolhode cones and curves for uniformly accelerated precession.

36. (page 55) Hamilton's disciple Peter Guthrie Tait, author of two books on quaternions, writes in an obituary of Hamilton [Tait 1866] that

> It is characteristic of Hamilton that he fancied he saw in the quaternion, with its scalar and vector elements, the one merely numerical, the other having reference to position in space, a realization of the Pythagorean *Tetractys*
>
> *παγὰν ἀενάου φύσεως ῥιζώματ᾽ ἔχουσαν*,
>
> as it is called in the *Carmen Aureum.*

The *Carmen Aureum* (Golden Verses) is a set of ethical precepts supposedly associated with Pythagoras. The quoted verse describes the Tetractys, a sacred Pythagorean symbol of the number four, as the "source of eternal nature."

Tait also quotes a sonnet that Hamilton wrote on the subject:

The Tetractys.

Or high Mathesis, with her charm severe,
  Of line and number, was our theme; and we
  Sought to behold her unborn progeny,
And thrones reserved in Truth's celestial sphere:
While views, before attained, became more clear;
  And how the One of Time, of Space the Three,
  Might, in the Chain of Symbol, girdled be:
And when my eager and reverted ear
Caught some faint echoes of an ancient strain,
  Some shadowy outlines of old thoughts sublime,
Gently he smiled to see, revived again,
  In later age, and occidental clime,
    A dimly traced Pythagorean lore,
    A westward floating, mystic dream of FOUR.

The sonnet, says Tait, "gives besides a good idea of his [Hamilton's] powers of poetical composition."

37. (page 56) *Drehstreckung.*

38. (page 59) *Wendestreckung.*

39. (page 59) The vector part of a quaternion has different transformation properties from what is today called a vector [Altmann 1989, p. 298].

40. (page 61) *Umklappungen.*

41. (page 61) See the supplementary note by Klein and Sommerfeld on page 198.

42. (page 62) This apparently defines the action of the quaternion Q when it acts as an operator on the right, as Klein and Sommerfeld use it on page 63. The derivation of page 63 can also be made by using Q as an operator on the left; Q then transforms the axis of V into the axis of v, and Q^{-1} transforms the axis of v into the axis of V.

43. (page 64) The geometric demonstration of equation (13) on page 64 is illustrated in Fig. 155. The axes of Q, v, and V remain fixed in space, while the sphere K, which contains a conical hole for visual aid, rotates and dilates.

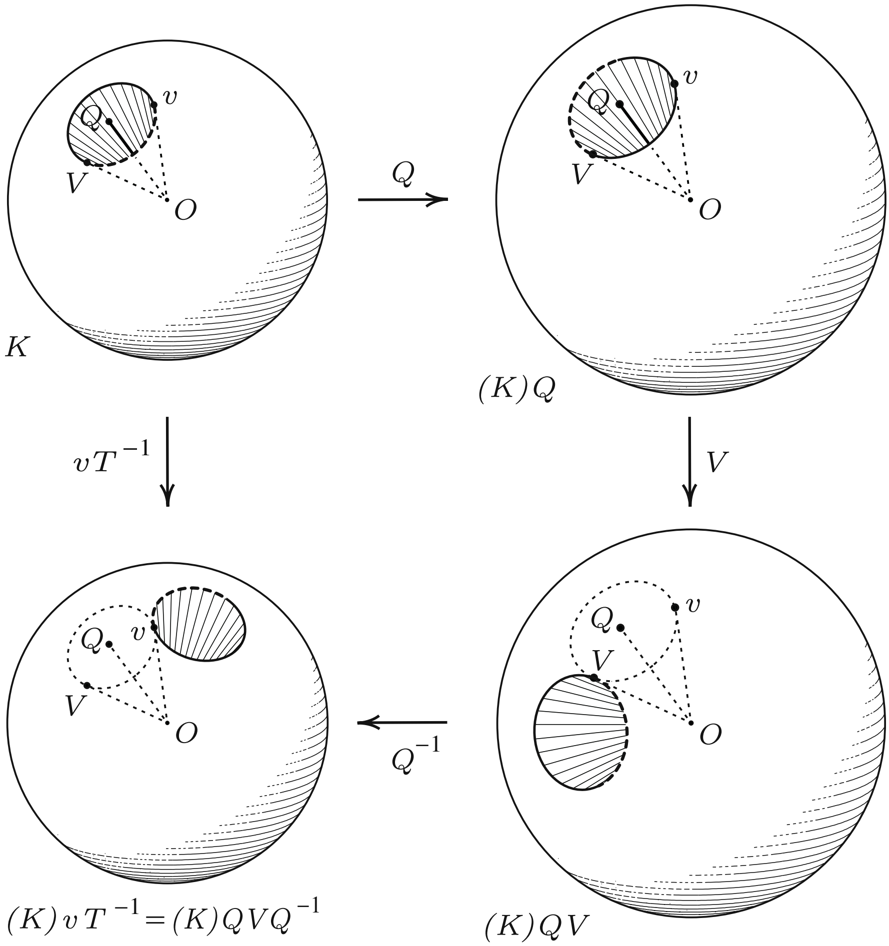


Fig. 155. Geometric demonstration of the quaternion formula $vT^{-1} = QVQ^{-1}$.

44. (page 64) See the supplementary note by Klein and Sommerfeld on page 199.

45. (page 68) See the supplementary note by Klein and Sommerfeld on page 204.

46. (page 69) A critical and historical account of the development of the principles of mechanics had been given in 1883 by Ernst Mach (1838–1916) in *Die Mechanik in ihrer Entwickelung; historisch-kritisch dargestellt.* Mach describes the anthropomorphic interpretation of force in his discussion of statics [Mach 1883, p. 78]:

> Let us, finally, cast a glance on the concept of force in statics. A force is a circumstance of which the consequence is motion. Several circumstances of this kind, each of which cause motion individually, can occur in combination without motion. Statics investigates only the necessary dependence of these circumstances on each other. Statics does not trouble itself further over the specific type of motion that is caused by a force. The motion causing circumstances that are best known to us are our own acts of will, our innervations. In the motions that we ourselves determine, as well as those to which we are forced by external circumstances, we always feel a pressure. Thus develops the habit of representing every motion causing circumstance as something related to an act of will and a pressure. The attempts to dispel this representation as subjective, animistic, and unscientific invariably fail. It cannot be useful to do violence to one's own natural thoughts and to condemn oneself voluntarily to distress. We will see that the present conception also plays a role in the foundation of dynamics.

Newton's concept of force is not as simple as Klein and Sommerfeld imply. In the *Definitions* section of the *Principia*, Newton distinguishes three different types of forces [Newton 1962, p. 2]:

> DEFINITION III
>
> *The* vis insita, *or innate force of matter, is a power of resisting, by which every body, as much as in it lies, continues in its present state, whether it be of rest, or of moving uniformly forwards in a right line.*
>
> DEFINITION IV
>
> *An impressed force is an action exerted upon a body, in order to change its state, either of rest, or of uniform motion in a right line.*

DEFINITION V

A centripetal force is that by which bodies are drawn or impelled, or in any way tend, towards a point as to a centre.

For Heinrich Hertz (1857–1894), the Newtonian conception of force and the corresponding laws of motion were logically untenable. Hertz gives the following criticism in *Die Prinzipien der Mechanik* [Hertz 1894, pp. 6–9]:

> We swing a stone on a cord in a circle; we thus, in a conscious measure, exert a force on the stone; this force constantly deflects the stone from its straight path, and if we change this force, the mass of the stone, and the length of the string, we find that the actual motion of the stone always follows, in fact, in agreement with Newton's second law. But now the third law demands an opposing force to the force that the hand exerts on the stone. To the question regarding this opposing force, the familiar answer is that the stone reacts upon the hand as a result of the centrifugal force, and that this centrifugal force is, in fact, exactly equal and opposite to the force that is exerted by us. Now, is this manner of expression permissible? Is what we now call the centrifugal force anything other than the inertia of the stone? May we, without destroying the clearness of our representation, take the action of inertia into account twice, namely once as mass, and secondly as force? In our laws of motion, force was the cause of motion, present before the motion. May we, without disordering our ideas, now suddenly speak of forces which first arise through motion, which are a consequence of motion? May we pretend that we had already stated something about this new type of force in our laws, that we can bestow by the name "force" the properties of a force as well? All these questions are obviously to be denied; there remains for us nothing more than to explain that the designation of the centrifugal force as a force is improper; that its name, like the name *vis viva*, is added as a historical tradition, more to be excused than justified. But what then remains of the demands of the third law, which requires a force that the inanimate stone exerts on the hand, and which must be satisfied by an actual force and not by a mere name?

I do not believe that these difficulties are evoked artificially or wantonly; they press upon us of themselves. Is not their origin to be traced back to the fundamental laws? The force of which the definition and the first two laws speak act upon a body in a single determined direction. The meaning of the third law is that forces always bind two bodies, and are directed just as well from the first to the second as from the second to the first. The representation of force that this third law assumes and awakens in us and the representation of the first two laws appear to me to differ by a small amount; but this small difference may very well produce the logical turbidity whose consequences burst forth in our example. We need not enter into the investigation of further examples. We can call upon general observations as witnesses for the justification of our doubt. A first such observation appears to me to be formed by the experience that it is very difficult to present an introduction to mechanics directly to thoughtful listeners without some embarrassment, without having to apologize here and there, without the desire to hurry rapidly from the preliminaries to examples that speak for themselves. I believe that Newton himself must have felt this embarrassment when he somewhat brutally defined mass as the product of volume and density. I believe that Thomson and Tait must have sympathized with him when they remarked that this is actually more a definition of density than of mass, and nevertheless satisfied themselves with this as the single definition of mass. I think Lagrange as well must have sensed this embarrassment, and the wish to proceed forward at any cost, when he prefaced his Mechanics abruptly with the explanation that a force is a cause that imparts "or strives to impart" a motion to a body, certainly not without feeling the logical harshness of such a definition. I draw additional evidence from the fact that we have many proofs by distinguished mathematicians of the elementary theorems of statics, of the theorem of the parallelogram of forces, of the theorem of virtual velocities, and so on, proofs that make the claim to be rigorous, and yet which in the judgment of other prominent mathematicians in no way satisfy this claim. In a logically complete science, in

> pure mathematics, a difference of opinion in such questions is absolutely unthinkable. The all too frequently heard assertions that the nature of force is still mysterious, that it is a major problem of physics to investigate the nature of force, and many similar statements, also appear to me as very incriminating evidence. In the same sense, electrical researchers are always besieged regarding the nature of electricity. Now, why is the nature of gold or the nature of velocity never questioned in this sense? Is the nature of gold better known than that of electricity, or the nature of velocity better known than that of force? Can we exhaustively reproduce the nature of anything by our representations, by our words? Certainly not. I believe that the difference is this: we associate with the symbols "velocity" and "gold" a large number of relations to other symbols, and we find no offensive contradictions among all these relations. This satisfies us, and we ask nothing further. But to the symbols "force" and "electricity," more relations have accumulated than can fully tolerate each other; we feel this dimly, demand an explanation, and express our confused desire in confused questions regarding the nature of force and electricity. But these questions are obviously mistaken with respect to what we want. We will be satisfied not by the recognition of new and more relations and connections, but rather by the removal of the contradictions among those at hand, and perhaps by reducing the number of the existing relations. If these painful contradictions are removed, the question of the nature of force is not answered, but the untormented mind will cease to ask itself this illegitimate question.

In *Die Prinzipien der Mechanik*, Hertz attempts to construct an image (*Bild*, mental representation) of mechanics that is based only on the concepts of space, time, and mass. It is astonishing to compare the abstract and formal development of Hertz, a physicist and a student of the physician and physicist Hermann Helmholtz (1821–1894), with the physical and intuitive presentation of the mathematicians Klein and Sommerfeld.

47. (page 70) The absolute system of measurement was proposed by Carl Friedrich Gauss (1777–1855) as part of his work with Wilhelm

Eduard Weber (1804–1891) on the measurement of the magnetic field of the Earth [Gauss 1863]. The fundamental physical dimensions in this system are mass, length, and time. The units of mass (gram) and length (centimeter) are defined by physical standards, and the unit of time (second) is defined in terms of the period of the Earth's rotation with respect to the stars. The unit of force (dyne) is the force that produces a unit increase in velocity when applied to a unit of mass for a unit of time, and the unit of work (erg) is the unit of force times the unit of distance. Experiments with falling bodies relate the gravitational force of the Earth at a particular location to the unit of force, and thus the unit of work to the work required to raise one gram by one centimeter.

48. (page 71) The impact force (*Stoſskraft*) used by Klein and Sommerfeld would be written in modern notation as

$$P(t) = [P]\delta(t - t_0),$$

where $[P]$ is the magnitude of the impact force, δ is the Dirac delta (or unit impulse) function, and t_0 is the time at which the impact force is applied.

49. (page 73) The cited paper by Ludwig Boltzmann (1844–1906) has been published in English under the title "On the indispensability of atomism in natural science" [Boltzmann 1974, pp. 41–53]. Boltzmann argues that it is impossible to conceive a continuum, and that the only rational scientific models are those in which a finite number of objects are subjected to a finite number of computational operations:

> Let us for example analyse the meaning of the classical instance of Fourier's equation for heat conduction. It expresses nothing else but a rule consisting of two parts:
>
> (1) Within a body (or, more generally, in a regular arrangement within a corresponding bounded three-dimensional manifold), imagine numerous small things (let us call them elementary particles, or, better still, elements or atoms in the most general sense), each of which has an arbitrary initial temperature. After a very short time has elapsed (or when the fourth variable increases by a small amount) let the temperature of each particle be the arithmetic mean of the initial temperatures of the particles that

had immediately surrounded it previously.* After a second equal lapse of time the process is repeated and so on.

(2) Imagine both the elementary particles and the increments of time becoming ever smaller and their number growing in corresponding proportion, and let them stop at those temperatures at which further diminution no longer noticeably affects the results.

Likewise, definite integrals that represent the solution of the differential equation can in general be calculated only by mechanical quadratures and thus again demand division into a finite number of parts.

Do not imagine that by means of the word continuum or the writing down of a differential equation, you have acquired a clear concept of the continuum. On closer scrutiny the differential equation is merely the expression for the fact that one must first imagine a finite number; this is the first prerequisite, only then is the number to grow until its further growth has no influence. What is the use of concealing the requirement of a large number of individuals now, when at the stage of explaining the differential equation one has used that requirement to define the value expressed by that number? My apologies for the somewhat banal expression, if I say that those who imagine that they have got rid of atomism by means of differential equations fail to see the wood for the trees. Explaining differential equations by complicated geometrical or other physical concepts would indeed help all the more to make the equation for heat conduction appear in the light of an analogy rather than of a direct description. In reality we cannot distinguish the neighbouring parts. However, a picture in which from the start we did not distinguish adjacent parts would be hazy; we could not apply to it the prescribed arithmetical operations.

If I then declare differential equations, or a formula containing definite integrals, to be the most appropriate picture, I surrender to an illusion if I imagine that I have thus banished atomistic conceptions from my mental pictures.

*Maxwell, *Treatise on Electricity*, 1873, Vol. 1, sec. 29; Mach, *Prinzipien der Wärmelehre*, Leipzig, 1893, p. 363. His writings on these matters have greatly helped in clarifying my own world view [Boltzmann's note].

Without them the concept of a limit is senseless; I merely add the further assertion that however much our means of observation might be refined, differences between facts and limiting values will never be observed.

50. (page 74) The quantity of motion is defined by Newton in Definition II of the *Principia*: *The quantity of motion is the measure of the same, arising from the velocity and quantity of matter conjointly* [Newton 1962, p. 1]. René Descartes (1596–1650) had used the same term and the same definition previously [Mach 1883, p. 254]. Newton's statement of the second law of motion (*The change of motion is proportional to the motive force impressed; and is made in the direction of the right line in which that force is impressed* [Newton 1962, p. 13]) refers to motion in the sense of the quantity of motion, and to force in the sense of the impact force used by Klein and Sommerfeld.

51. (page 74) We have used *vis viva* as a translation of the original German *lebendige Kraft*. The concept of the *vis viva* was introduced by Gottfried Wilhelm Leibniz (1646–1716), who defined it as mv^2. Gaspard-Gustave de Coriolis (1792–1843) later redefined the *vis viva* as $\frac{1}{2}mv^2$ [Mach 1883, p. 254]. The modern term kinetic energy is introduced by Klein and Sommerfeld on page 117.

52. (page 81) The Lagrange equations of the second kind, which are now usually known simply as the Lagrange equations, look more familiar if equations (16) on page 79 are combined with equations (18) on page 81 to give, for example,

$$\frac{d}{dt}\frac{\partial T}{\partial \varphi'} - \frac{\partial T}{\partial \varphi} = \Phi.$$

The Lagrange equations of the first kind involve the use of Lagrange multipliers to account for the effects of kinematic constraints; these equations are discussed by Klein and Sommerfeld in Ch. 3, §3.

53. (page 81) In particle mechanics, Thomson and Tait [Thomson 1883] and Maxwell [Maxwell 1991] use both the term momentum and the term impulse. The momentum of a particle is defined to be the product of the mass m and the velocity $\mathbf{v}$, and the impulse (or, for Thomson and Tait, impact) of a continuously applied force is defined to be the time integral of the force. Maxwell states Newton's second law explicitly in terms of impulse and momentum: *The change of momentum of a body is numerically equal to the impulse which produces*

it, and is in the same direction. The illuminating conception of Klein and Sommerfeld is the interpretation of the total momentum $m\mathbf{v}$ as the impact force that is capable of generating the actual motion instantaneously, and in the instantaneous position, from rest.

For problems involving the motion of a solid body in an ideal fluid, Thomson defined the impulse as the impulsive wrench applied to the solid that might have generated the instantaneous motion of the solid and the fluid [Lamb 1945, p. 161]. The mention of Maxwell's attempt to use energy as the basis of the general equations of mechanics may also refer to the discussion of the Lagrange equations for general dynamic systems in Vol. II, Part IV, Ch. V of the *Treatise on Electricity and Magnetism* [Maxwell 1873, pp. 185–194; Maxwell 1991, pp. 123–136].

54. (page 82) The axiom that a force on a rigid body can be displaced arbitrarily along its line of action is sometimes called the principle of transmissibility of force. Poinsot considers this principle as a corollary to the seemingly simpler axiom that two equal and opposing forces that act at different points on the same line are in equilibrium. It is evident to Poinsot that two such forces are in equilibrium, "for there is no reason for motion to be produced (*naisse*) to one side rather than the other" [Poinsot 1811, p. 13].

55. (page 82) The law of the lever may be derived from the parallelogram law and the transmissibility of force in the following manner. Suppose the oppositely directed forces P and Q are applied to the rigid line segment AB, as shown in Fig. 156. The equal, opposing and collinear forces M and N may be added without effect. Using the parallelogram law, the forces P and M are equivalent to the force S, and the forces Q and N are equivalent to the force T. The forces S and T are equivalent, according to the principle of transmissibility, to the forces S' and T' at the point D, the intersection of the lines of action of S and T. When S' and T' are composed according to the parallelogram law, the opposing components M and N cancel, leaving a resultant R' that has magnitude $Q - P$ and is parallel to P and Q. Again using the principle of transmissibility, the two original forces P and Q are therefore equivalent to the single force R at the point C on the line through AB. The position of the point C and the magnitude of R are independent of the arbitrarily chosen common magnitude of M and N. By similar triangles,

$$\frac{P}{M} = \frac{CD}{AC}, \quad \frac{Q}{N} = \frac{CD}{BC}.$$

Since M and N are equal,

$$P \cdot AC = Q \cdot BC,$$

which is the familiar law of the lever.

As the forces P and Q become equal, the magnitude of R approaches zero, while the lines AD and BD become parallel, so that the points C and D are removed to infinity.

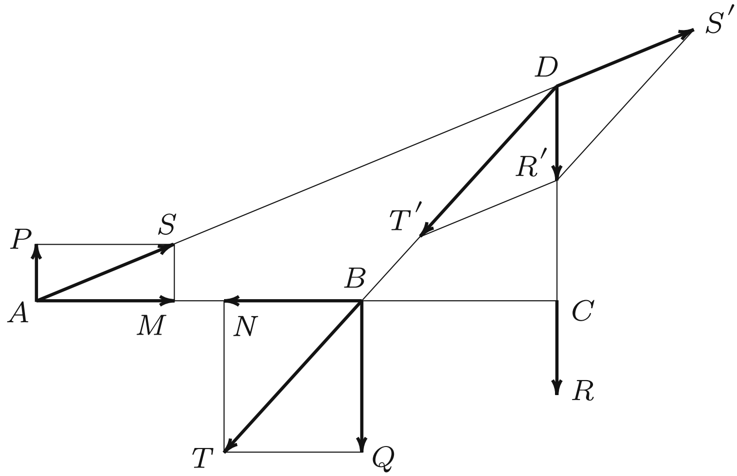


Fig. 156. Derivation of the lever law for oppositely directed forces P and Q applied to the rigid line segment AB.

Poinsot does not use the parallelogram law in his derivation of the law of the lever [Poinsot 1811, pp. 21–24]. He uses, instead, a symmetry argument of the type originally due to Archimedes [Mach 1883, pp. 8–19]. Poinsot then derives the parallelogram law on the basis of the law of the lever and the transmissibility of force [Poinsot 1811, pp. 31–35]. For Poinsot, who begins directly with the statics of rigid bodies, the law of the lever is simpler and more fundamental than the parallelogram law; he remarks that the law of the lever was known in ancient times, while the parallelogram law dates only from the time of Galileo [Poinsot 1811, p. xij].

56. (page 83) *linienflüchtiger.*

57. (page 88) The *Lehrbuch der Statik* by August Ferdinand Moebius (1790–1868) is included in Vol. 3 of his complete works [Moebius 1885, Vol. 3, pp. 1–497]; the editor of this volume is Felix Klein. In the

first edition of *The Theory of Screws*, Robert Stawell Ball (1840–1913) gives a description of Moebius's work [Ball 1876, pp. 178–179]:

> This book is, we learn from the preface, one of the numerous productions to which the labours of Poinsot has given rise. The first part, pp. 1–355, discusses the laws of equilibrium of forces, which act upon a single rigid body. The second part, pp. 1–313, discusses the equilibrium of forces acting upon several rigid bodies connected together. The characteristic feature of the book is its great generality. I here enunciate some of the principal theorems.
>
> If a number of forces acting upon a free rigid body be in equilibrium, and if a straight line of arbitrary length and position be assumed, then the algebraic sum of the tetrahedra, of which the straight line and each of the forces in succession are pairs of opposite edges, is equal to zero (p. 94).
>
> If four forces are in equilibrium they must be generators of the same hyperboloid (p. 177).
>
> If the *lines of action* of five forces be given, then a certain plane S through any point P is determined. If the five forces can be equilibrated by one force through P, then this one force must lie in S (p. 180).
>
> To adopt the notation of Professor Cayley, we denote by 12 the perpendicular distance between two lines 1, 2, multiplied into the sine of the angle between them (Comptes Rendus, t. lxi., pp. 829–830). Moebius shows (p. 189) that if forces along four lines 1, 2, 3, 4 equilibrate, the intensities of these forces are proportional to
>
> $$\sqrt{23\,.\,24\,.\,14},\quad \sqrt{13\,.\,14\,.\,34},\quad \sqrt{12\,.\,14\,.\,24},\quad \sqrt{12\,.\,13\,.\,23}.$$
>
> It is also shown that the product of the forces on 1 and 2, multiplied by 12 equals the product of the forces on 3 and 4 multiplied by 34. He hence deduces Chasles' theorem (Liouville's Journal, t. xii., p. 222), that the volume of the tetrahedron formed by two of the forces is equal to that formed by the remaining two.

It was a great surprise for the translators to find that so many prominent nineteenth-century mathematicians devoted their attention to the statics of rigid bodies, and developed it to such an extent. The subject

is now neglected entirely in physics and mathematics, and covered only superficially in engineering curricula.

58. (page 91) The mechanical concept of work was first defined and used in its modern sense by Coriolis in his treatise *Du calcul de l'effet des machines* [Coriolis 1829]. Coriolis was professor of mechanics in the *École Centrale des Artes et Manufactures*, the *École des Ponts et Chaussées*, and the *École Polytechnique* in Paris.

59. (page 102) If the rotation vector is OP, where O is the center of the ellipsoid and P is a point on the ellipsoid, then the conjugate plane to the rotation vector is the plane through O that is parallel to the tangent plane at P.

60. (page 106) If the corners of the square have mass m_0 and the side length of the square is L, then

$$m = \left(\frac{P}{g}\right)^2 \frac{1}{m_0 L^2}, \qquad E = \frac{m_0 L^2 g}{P}.$$

61. (page 112) The area theorems, which together constitute what is now called the angular momentum principle, receive their names from the fact that $xy' - yx'$, for example, is the time rate of change of the area described in the xy-plane by the projection of the position vector of a particle with coordinates (x, y, z). Kepler's second law of planetary motion is a familiar example of the relation between angular momentum and the time rate of change of area described by the radius vector of a particle.

62. (page 113) Robert Baldwin Hayward (1829–1903) studied mathematics in University College, London, and St. John's College, Cambridge. In the paper cited by Klein and Sommerfeld [Hayward 1856], Hayward derives and gives several applications of the formula that is now written in vector form as

$$\frac{d\mathbf{u}}{dt} = \left(\frac{\partial \mathbf{u}}{\partial t}\right)_{\mathrm{rel}} + \boldsymbol{\omega} \times \mathbf{u},$$

where $d\mathbf{u}/dt$ is the time derivative of the vector $\mathbf{u}$ with respect to a fixed spatial frame, $(\partial \mathbf{u}/\partial t)_{\mathrm{rel}}$ is the time derivative of $\mathbf{u}$ with respect to the moving frame, and $\boldsymbol{\omega}$ is the angular velocity of the moving frame with respect to the fixed frame.

In 1859 Hayward became a mathematics master in the Harrow School in London, where he remained for thirty-four years.

63. (page 116) The increments dL, dM, dN in the preceding expression are to be interpreted as changes with respect to the body frame.

64. (page 122) The rolling ellipsoid, the polhode curve, and the herpolhode curve are illustrated in Fig. 157 for the case corresponding to the parameters defined on page 125.

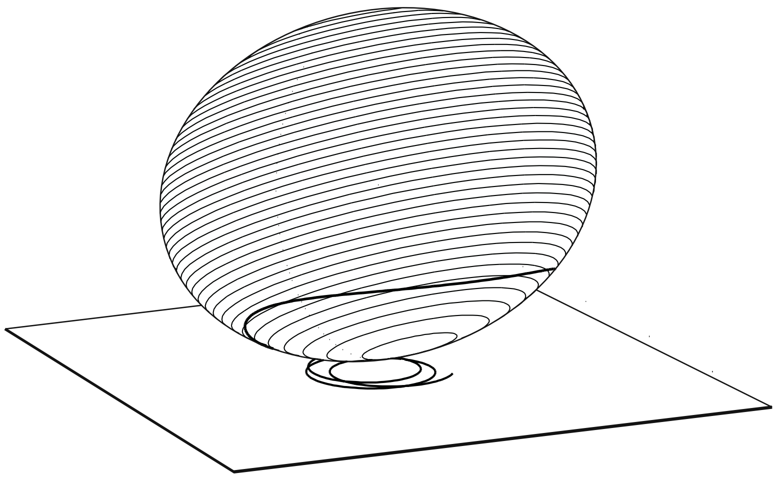


Fig. 157. Ellipsoid rolling on a plane, with the polhode curve on the ellipsoid and the herpolhode curve coiled on the plane.

65. (page 122) James Joseph Sylvester (1814–1897) was an English mathematician who taught from 1877 to 1884 in the Johns Hopkins University, where he founded the *American Journal of Mathematics*. (An earlier appointment in the University of Virginia ended after Sylvester struck an unruly student with a sword.) In his paper on the motion of a rigid body [Sylvester 1866], Sylvester begins with the observation that the rolling ellipsoid in Fig. 157 would be unaffected by the addition of a second plane parallel to the illustrated plane, and also tangent to the ellipsoid; the ellipsoid would then "roll between the two planes instead of rolling upon the under one alone." Sylvester then imagines that the upper portion of the rolling ellipsoid is modified ("pared away") so that its upper surface assumes the form of an ellipsoid *confocal* with the lower one; if the squared semiaxes of the lower ellipsoid are a^2, b^2, c^2, then those of the upper ellipsoid are $a^2-\lambda$, $b^2-\lambda$,

$c^2 - \lambda$. The parameter λ can be chosen in the range $0 < \lambda \leq c^2$, where c is the minimum semiaxis. For $\lambda = c^2$, the upper ellipsoid "degenerates into a curve or hoop." Sylvester shows that if a parallel tangent plane is placed on the modified ellipsoid, as illustrated in Fig. 158, friction will cause the upper plane to rotate with *constant angular velocity* about the normal through the center of the ellipsoid as the lower surface rolls according to the Poinsot representation on the lower fixed plane. If, as Sylvester assumes, the lower ellipsoid in Fig. 158 is the ellipsoid of inertia, then the angular velocity of upper plane is $L\lambda$, where L is the magnitude of the angular momentum of the ellipsoid. If, conversely, the upper plane is driven externally with the constant angular velocity $L\lambda$, friction will cause the lower ellipsoid to roll in exactly the manner required by the Poinsot representation.

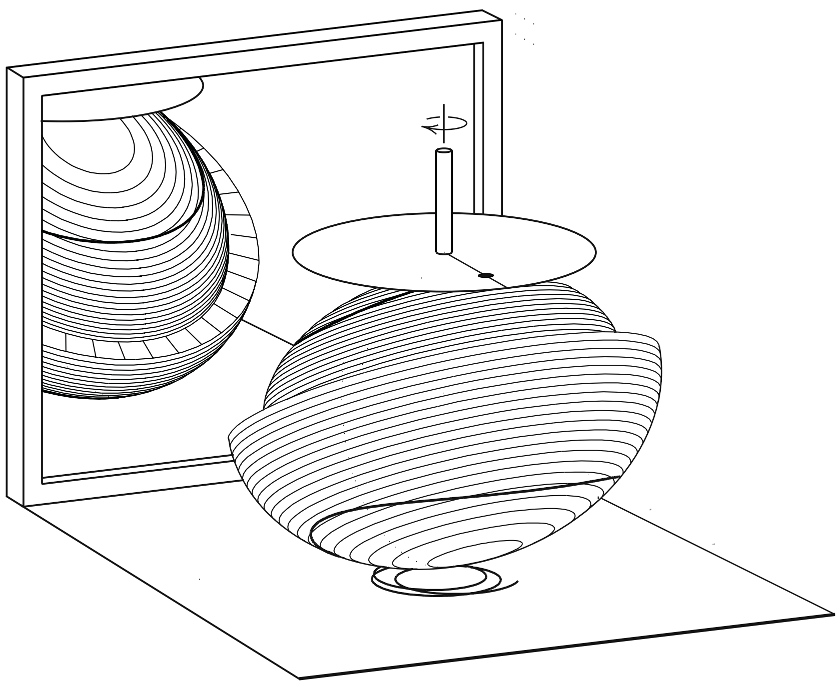
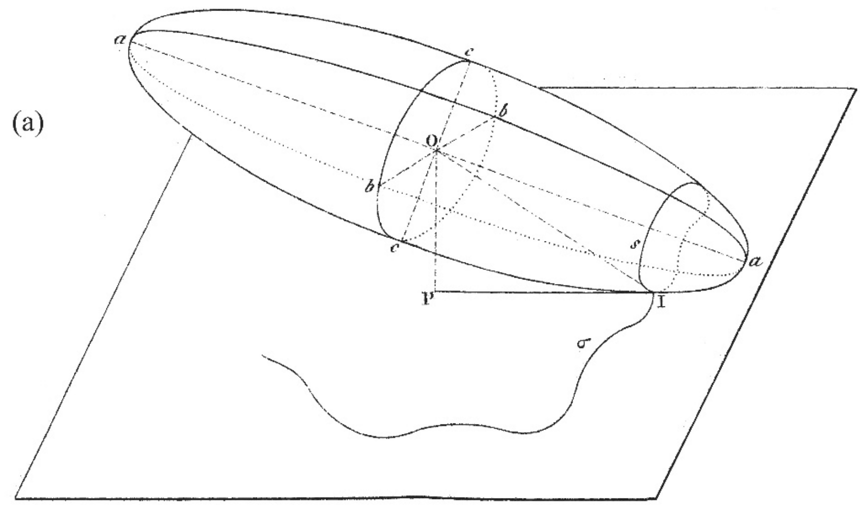


Fig. 158. Sylvester's construction for observing or regulating the time dependence of the rolling ellipsoid motion.

66. (page 125) Wilhelm Philipp Hess (1858–1937) published two long papers on rigid body motion in the *Mathematischen Annalen* [Hess 1882; Hess 1887]; these papers give many specific examples of polhode

and herpolhode curves for the top. In 1888 Hess became professor of physics, mathematics, and astronomy in the Bamberg Lyceum, where he taught until 1925.

Poinsot's illustration of the rolling ellipsoid with the polhode curve s and the herpolhode curve σ is shown in Fig. 159(a). An accurate plot of a herpolhode curve corresponding to the form of Poinsot's illustration is given in Fig. 159(b). As proven by Hess, the herpolhode curve in Fig. 159(b) has no inflection points. The curve is always concave toward the center of the dashed concentric circles, except at the points of tangency with the inner circle, where the concavity vanishes. The error in Poinsot's illustration, if there is one, is very minor.



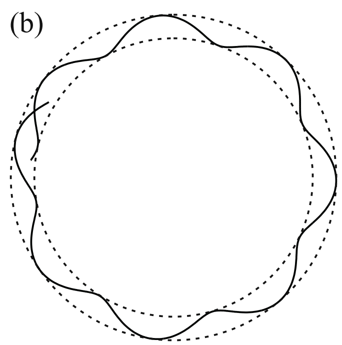


Fig. 159. Herpolhode curves produced by a rolling ellipsoid.
(a) Poinsot's illustration [Poinsot 1851, plate II];
(b) computational reproduction of Poinsot's herpolhode curve.

67. (page 132) In general, the herpolhode curve assumes a spiral form when $G^2 = 2hB$ and the initial position of the angular velocity does not coincide with the intermediate principal axis. The rolling ellipsoid for such a case is illustrated in Fig. 160. This special case is treated in detail by Poinsot, who gives the following derivation of an explicit formula for the herpolhode curve [Poinsot 1851, pp. 102–118].

According to equations (3) and (4) on page 124, a point (p, q, r) of the polhode curve lies on the two ellipsoids

$$\frac{p^2}{a^2} + \frac{q^2}{b^2} + \frac{r^2}{c^2} = 1 \tag{2}$$

and

$$\frac{p^2}{a^4} + \frac{q^2}{b^4} + \frac{r^2}{c^4} = \frac{1}{\varrho^2}, \tag{3}$$

where

$$a = \sqrt{2h/A}, \quad b = \sqrt{2h/B}, \quad c = \sqrt{2h/C}$$

are the principal semiaxes of the rolling ellipsoid, and $\varrho = 2h/G$ is the constant distance OP from the fixed center of the ellipsoid to the

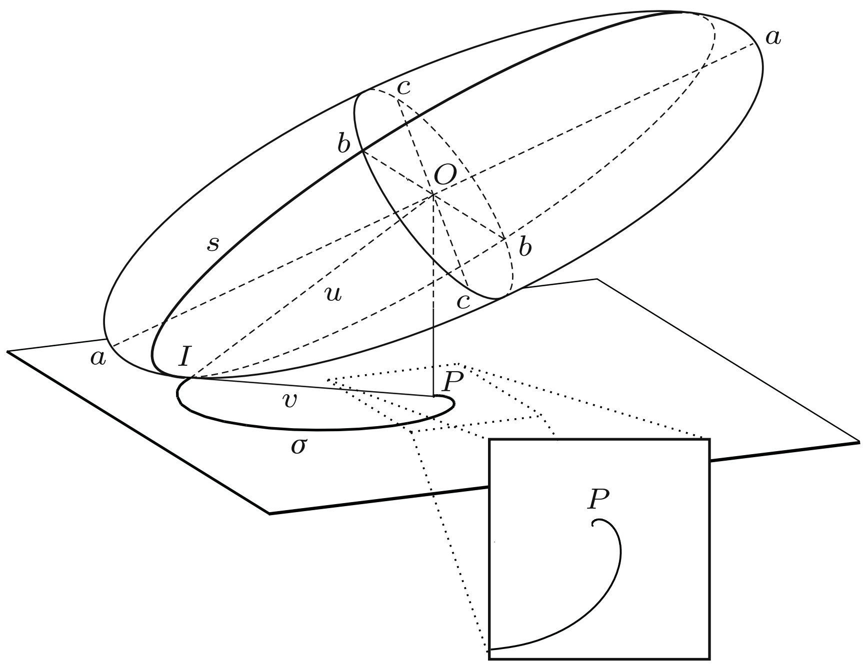


Fig. 160. Rolling ellipsoid for the special case in which the herpolhode curve assumes a spiral form.

horizontal plane. The line segment OI from the center of the ellipsoid to a point on the polhode curve has length

$$u = \sqrt{p^2 + q^2 + r^2}. \tag{4}$$

For the special case $G^2 = 2hB$,

$$\varrho = \frac{2h}{G} = \sqrt{\frac{2h}{B}} = b,$$

and equations (2) and (3) give

$$p^2 \left(\frac{a^2 - b^2}{a^4} \right) = r^2 \left(\frac{b^2 - c^2}{c^4} \right),$$

which defines the two planes containing the intersection of the ellipsoids (2) and (3). The intersection consists of two ellipses with semimajor axis β, where

$$\beta^2 = c^2 + a^2 - \frac{c^2 a^2}{b^2}.$$

One such ellipse s is illustrated in Fig. 160.

Using equations (2), (3), and (4), p, q, and r can be expressed in terms of u as

$$\begin{aligned}
p^2 &= \frac{-a^4}{(a^2 - b^2)(c^2 - a^2)} \left[u^2 - b^2 \right], \\
q^2 &= \frac{-b^4}{(b^2 - c^2)(a^2 - b^2)} \left[u^2 - \beta^2 \right], \\
r^2 &= \frac{-c^4}{(c^2 - a^2)(b^2 - c^2)} \left[u^2 - b^2 \right].
\end{aligned}$$

The differential arc length ds of the polhode curve is thus

$$\begin{aligned}
ds &= \sqrt{(dp)^2 + (dq)^2 + (dr)^2} \\
&= u\, du \, \frac{\sqrt{u^2 - (b^2 + \beta^2)}}{\sqrt{(u^2 - b^2)(u^2 - \beta^2)}}.
\end{aligned}$$

Since the polhode curve rolls on the herpolhode curve, the differential arc length $d\sigma$ of the herpolhode curve is $d\sigma = ds$. Also, from Fig. 160,

$$u^2 = v^2 + (OP)^2 = v^2 + b^2,$$

where v is the radius of the herpolhode curve, so that $u\,du = v\,dv$. If v, θ are taken as the polar coordinates of the planar herpolhode curve, then

$$d\sigma = \sqrt{v^2(d\theta)^2 + (dv)^2},$$

which gives the differential equation for the herpolhode curve in the form

$$d\theta = b\,\frac{dv}{v\sqrt{(\beta^2 - b^2) - v^2}}.$$

This equation can be integrated to give

$$\theta = \frac{b}{n}\log\left(\frac{v}{n + \sqrt{n^2 - v^2}}\right),$$

where it has been assumed that $\theta = 0$ when v attains its maximum value $n = \sqrt{\beta^2 - b^2}$. Solving for the radius v gives

$$v = \frac{2n}{e^{n\theta/b} + e^{-n\theta/b}},$$

which is the polar form of the equation for the herpolhode spiral.

Poinsot goes further with this special case, and gives the angle θ and the radius v in terms of the time t [Poinsot 1851, pp. 119–129, 289–301]. He shows that the herpolhode curve approaches the spiral center P only asymptotically as $t \to \infty$; as the contact point between the rolling ellipsoid and the horizontal plane approaches the end of the intermediate principal axis bb, the ellipsoid spins about this more and more nearly vertical axis.

68. (page 134) John Perry (1850–1920) studied engineering with Prof. James Thomson (1822–1892) in Queen's College, Belfast, and in 1874 became an assistant to Prof. Sir William Thomson (Lord Kelvin) in the University of Glasgow. Perry was professor of engineering in the Imperial College of Engineering in Tokyo, Japan, from 1875 to 1878, professor of mechanical engineering in the City and Guilds of London Technical College in Finsbury from 1878 to 1896, and professor of mechanics and physics in the Royal College of Science in London from 1896 to 1913. His book *Spinning Tops* [Perry 1890] is based on one of the many popular lectures given by Perry on scientific and technical subjects.

69. (page 139) The moving frame XYZ and the vectors i, J, R, di, and dJ are illustrated in Fig. 161. The moving coordinate frame XYZ rotates with angular velocity R with respect to a fixed spatial frame. In the moving frame, the displacement vector of the fixed spatial point P in the time interval dt is therefore $-(R \times J)\,dt = V(J, R)\,dt$. See also the supplementary note by Klein and Sommerfeld on page 204.

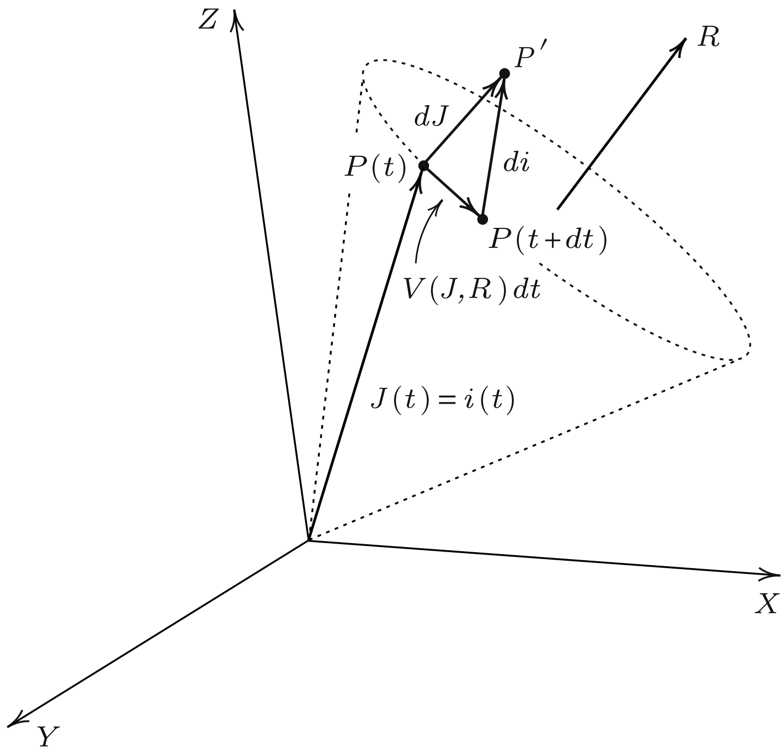


Fig. 161. Moving frame XYZ, with rotation vector R and vectors i, J, di, and dJ.

70. (page 142) A paper by Leonhard Euler (1707–1783) with the title *"Du mouvement de rotation de corps solides autour d'un axe variable"* appears in the *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin* for the year 1758 [Euler 1758]. In this paper, Euler derives the Euler equations, reduces their solution to quadratures for the case of torque-free motion, and gives the explicit solution for the torque-free motion of a body with two equal principal moments of

inertia. See also the supplementary note by Klein and Sommerfeld on page 205 and translators' note 107.

71. (page 142) Tait's work is included in his *Scientific Papers* [Tait 1898, Vol. I, pp. 86–127]. Although he professes great admiration for Poinsot's geometric solution of rigid body motion, Tait claims that he and his quaternions can do just as well: "... let the reader bear in mind that a *quaternion* equation is quite as suggestively intelligible, to those who understand it, as any geometrical diagram can possibly be." If one takes the trouble to translate Tait's quaternion equations into vector notation, parts of Tait's paper are remarkably close to the modern treatment of rigid body motion.

72. (page 151) The derivation and integration of the dynamic equations for the heavy spherical top in terms of α, β, γ, δ are discussed by Klein and Sommerfeld in Vol. II, Chapter VI, §9.

73. (page 151) Carl Gustav Jacob Jacobi (1804–1851) was professor of mathematics in the University of Königsberg. His paper *Sur la rotation d'un corp* [Jacobi 1850] is an extract from a letter addressed to the Academy of Sciences in Paris. Jacobi begins his fifty-eight page paper thus: "The problem of the rotation of an arbitrary solid body, which is not subject to any accelerating force, admits of solution by formulas so elegant and so perfect, that I cannot prevent myself from communicating them to your illustrious academy."

74. (page 151) The monograph *Sur quelques applications des fonctions elliptiques* by Charles Hermite (1822–1901) was published in 1885 [Hermite 1885]. The reprint of the monograph included in the collected works of Hermite [Hermite 1905] lists the following references to earlier publication of this work: *Comptes rendus de l'Académie des Sciences*, t. LXXXV, 1877, p. 689, 728, 821, 870, 984, 1085, 1185; t. LXXXVI, 1878, p. 271, 422, 622, 777, 850; t. LXXXIX, 1879, p. 1001, 1092; t. XC, 1880, p. 106, 201, 478, 643, 761; t. XCIII, 1881, p. 920, 1098; t. XCIV, 1882, p. 186, 372, 477, 594, 753. The application of Hermite's work on elliptic functions to the motion of the top is discussed by Klein and Sommerfeld in Vol. II, Chapter VI, §9.

75. (page 159) The use of supernumerary impulse coordinates is discussed by Klein and Sommerfeld in Vol. II, Chapter VI, §9.

76. (page 161) The Norwegian mathematician Marius Sophus Lie (1842–1899) was an early associate and collaborator of Felix Klein. Klein acknowledges his debt to Lie in the Preface to the *Vorlesungen über das Ikosaeder* [Klein 1884, p. iv]:

> My obligation to Herr Lie goes back to the years 1869–70, when we concluded our student-time in Berlin and Paris in intimate intercourse with one another. We then conjointly conceived the idea of considering, in general, geometric or analytic forms that are transformed into themselves through *groups of changes*. This idea has remained definitive for our double-sided later works, as far apart as they appear to lie. While I fixed my own attention primarily on *discrete* operations, and was therefore led, in particular, to the investigation of regular solids and their relation to the the theory of equations, Herr Lie attacked from the beginning the more difficult theory of *continuous* transformation groups, and thus *differential equations*.

77. (page 161) Karl Otto Heinrich Liebmann (1874–1939) received the Ph.D. degree in mathematics from the University of Jena in 1895, and became an assistant in the mathematical model collection in Göttingen in 1897. He was professor of mathematics in the University of Heidelberg from 1920 to 1935. Liebmann's paper in Vol. 50 of the *Mathematische Annalen* [Liebmann 1898] is a continuation of work by Tullio Levi-Cività (1873–1941) on the search for integrable cases of the dynamic equations for the motion of the top. Liebmann distinguishes twenty-five different types of tops, sixteen of which allow the integration of their dynamic equations by quadrature.

78. (page 167) The form of D'Alembert's principle used by Klein and Sommerfeld is presented and discussed by Routh [Routh 1960, pp. 45–51]. Routh also gives a short excerpt from the *Traité de dynamique* by Jean le Rond D'Alembert (1717–1783). D'Alembert states his principle in the four pages of the first chapter of the second part of the *Traité*, which we translate from the original French as follows [D'Alembert 1990, pp. 72–74]:

SECOND PART

General Principle to find the Motion of several Bodies which act on one another in an arbitrary manner; with several applications of this Principle (∗).

FIRST CHAPTER

Exposition of the Principle

Bodies act on one another in only three different manners which are known to us: by immediate impulsion, as in an ordinary collision; by means of some bodies interposed between them, & which are attached to them; or finally by virtue of a reciprocal attraction, as do the Sun & the Planets in the Newtonian system. The effects of this last type of action having been sufficiently examined, I will confine myself here to treat of the motion of bodies which collide in an arbitrary manner, or those which pull on each other by ropes or inflexible rods. I limit myself the more willingly to this subject, since the greatest Geometers have solved until now (1742) only a very small number of Problems of

(∗) This principle & the greater part of the following Problems are contained in a Memoir which I have read to the Académie toward the end of 1742, even though the first Edition of this Treatise has not appeared until 1743. On the same day in which I started the reading of my Memoir, M. Clairaut presented one which has for its title *On some Principles which facilitate the solution of a great number of Problems in Dynamics*; this Memoir, printed in the Volume of 1742, has been read after mine, with which it has, moreover, nothing in common.

this type, & since I hope, by the general Method which I will give, to enable all those who are familiar with the calculations and principles of Mechanics to solve the most difficult problems of this kind.

Definition

In the following, I will call the *Motion* [*Mouvement*] of a body the speed [*vitesse*] of this same body, considered with regard to its direction; & I will understand by the *quantity of Motion* [*quantité de Mouvement*], as usual, the product of the mass and the speed.

GENERAL PROBLEM

60. *Let there be given a system of bodies disposed with respect to each other in an arbitrary manner; & suppose that one impresses to each of these Bodies a particular Motion, which it is not able to follow due to the action of the other Bodies; find the Motion that each Body must acquire.*

Solution

Let A, B, C, &c. be the bodies which compose the system, & let us suppose that one has impressed to these bodies the motions a, b, c, &c., which are forced, due to their [the bodies] mutual action, to change into the motions a, b, c, &c. It is clear that one can regard the motion a impressed on body A as composed of the motion a, which it has acquired, and another motion α; one can, in the same way, regard the motions b, c, &c. as composed of the motions b, β; c, $\varkappa$; &c., from which it follows that the motion of the interacting bodies A, B, C, &c. would have been the same if, instead of giving them the impulsions a, b, c, one would have given them at the same time the double impulsions a, α; b, β; c, $\varkappa$, &c. Now by assumption the bodies A, B, C, &c. have acquired the motions a, b, c; &c. Thus the motions α, β, $\varkappa$ &c. must be such that they do not disturb in any way the motions a, b, c, &c.; that is to say, that if

the bodies had received only the motions α, ϵ, ϰ &c., these motions would have destroyed themselves mutually, & the system would have remained at rest.

From this there results the following principle for finding the motion of several bodies which act upon one another. *Decompose the motions a, b, c,* &c. *impressed to each body, each into two others* a, α; b, ϵ; c, ϰ; &c., *which are such that if one had impressed to the bodies only the motions* a, b, c &c., *they could have conserved their motions without interfering with one another*; & *that if one had impressed only the motions* α, ϵ, ϰ, &c., *the system would have remained at rest;* it is clear that a, b, c will be the motions that the bodies will acquire in virtue of their [the motions *a*, *b*, *c*, &c.] action. *This is what was to be found.*

COROLLARY

When one of the impressed motions is = 0, it is evident that the motions in which one decomposes it are equal & contrary motions. For example, if *a* is = 0, one will have the motion α = & in direction contrary to the motion a: in effect, *a* is in all cases the diagonal of a parallelogram in which a & α are the sides; now when the diagonal is = 0, the sides are equal & oppositely directed. Thus &c.

D'Alembert's own use of his principle is discussed in two interesting papers by Craig Fraser [Fraser 1985a; Fraser 1985b].

79. (page 184) The paper cited by Klein and Sommerfeld is the second of two papers [Coriolis 1832; Coriolis 1835] in which Coriolis derives the dynamic equations for a general mechanical system with respect to an arbitrarily moving reference frame.

80. (page 189) Karl Ernst Ritter von Baer, Edler von Huthorn (1792–1876) was professor of anatomy and zoology in the University of Königsberg. He moved to Russia in 1834, and became academician for comparative anatomy and physiology in the Academy of Sciences in St. Petersburg. He was a founder and the first president of the Russian Geographical Society. Baer's law (written by Klein and Sommerfeld as *Beer'schen Gesetz*) states that a northward flowing river in the Northern Hemisphere will erode its bank most strongly on the right-hand side. Albert Einstein (1879–1955) published a paper on Baer's law in 1926

[Einstein 1926]. Einstein argues that the Coriolis force on a northward flowing river causes a nonuniform circulatory flow in the cross section of the river, and that this flow is strongest and most erosive on the right-hand side.

In the English translation of Einstein's paper [Einstein 1982, pp. 249–253], the circulatory flow in Fig. 2 is illustrated incorrectly. In Fig. 2 of the original, the circulatory flow always proceeds clockwise.

81. (page 189) Georg August Dietrich Ritter (1826–1908) became the first director of the *Institut für Allgemeine Mechanik* at the *Technische Hochschule* in Aachen in 1870. He was succeeded as director in 1900 by Arnold Sommerfeld. The institute is active today in the field of nonlinear solid mechanics.

82. (page 190) Christoph Hendrik Diederik Buys-Ballot (1817–1890) was professor of mathematics and physics in the University of Utrecht. The Buys-Ballot law states that wind flows counterclockwise (as seen from above) around a low pressure region in the Northern Hemisphere, and clockwise around a low pressure region in the Southern Hemisphere.

83. (page 190) According to the article cited by Klein and Sommerfeld [Jouffret 1874], Esprit Pascal Jouffret (1837–19??) was *capitaine d'artillierie* and professor in *l'École d'applications de l'artillerie et du génie à Fontainebleau.* His later textbook on four-dimensional geometry [Jouffret 1903] is said to have influenced the development of cubist art [Robbin 2006].

84. (page 190) August Christian Wilhelm Hermann Scheffler (1820–1903) was a physicist and engineer who held government positions as building inspector, surveyor, and finance secretary in Brunswick. He is the author of at least twenty-six books, on topics including mathematics, vision, steam boiler explosions, tax policy, and building construction. His paper *Imaginäre Arbeit, eine Wirkung der Centrifugal- und Gyralkraft* was published in the *Zeitschrift für Mathematik* [Scheffler 1866]. Scheffler defines the imaginary work on a particle in terms of the change in the direction of the particle motion, just as the conventional work on a particle can be defined in terms of the change in the speed of the particle motion. Scheffler's extension of the concept of work has apparently found very little use in mechanics.

85. (page 191) The resistance couple R for regular precession of the hand-top is illustrated in Fig. 162. The figure axis of the top is F and the axis of precession is V.

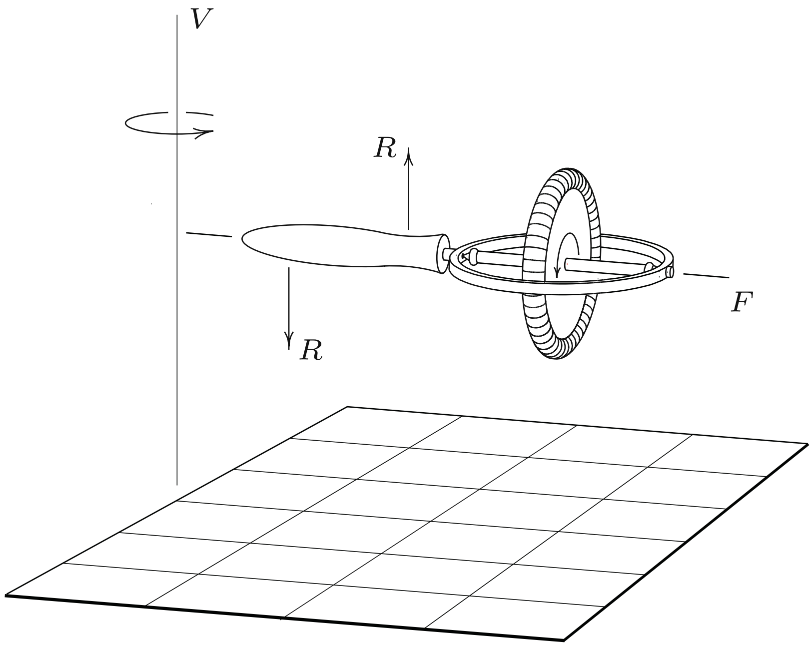


Fig. 162. Resistance couple R for a hand-top with figure axis F and axis of precession V.

86. (page 192) The motion of the balance illustrates the basic principle of the single-axis rate gyroscope, an important instrument in inertial navigation systems [Arnold 1961, Chs. 13, 14].

87. (page 192) Klein and Sommerfeld never return to Kelvin's kinematic theory of matter. A mechanical model similar to the device described by Klein and Sommerfeld is used by Kelvin in his paper "On a gyrostatic adynamic constitution for 'ether'" [Thomson 1890, pp. 466–472]. This work of Kelvin is cited in Sommerfeld's 1892 paper *"Mechanische Darstellung der electromagnetischen Erscheinungen in ruhenden Köpern"* [Sommerfeld 1892].

88. (page 195) The configurations of the hand-top before and after the handle rotation v are illustrated in Figs. 163(a) and 163(b), respectively.

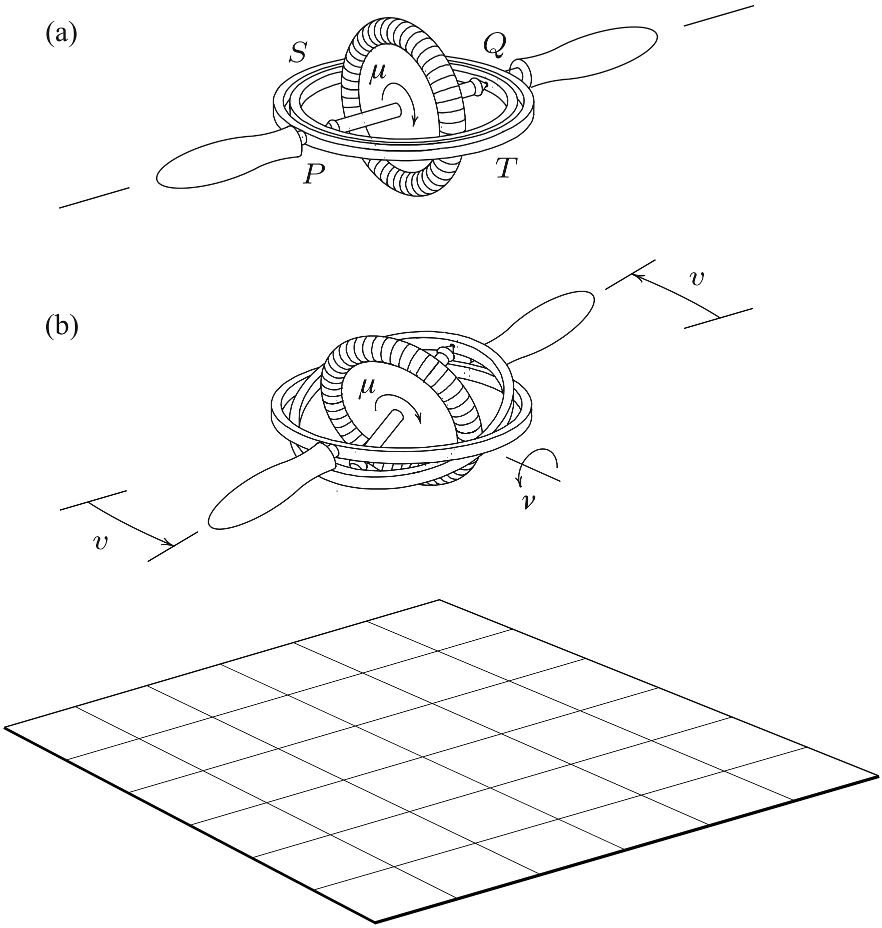


Fig. 163 Hand-top with two degrees of freedom.
(a) Initial configuration.
(b) Configuration after the rotation of the handle through the angle v.

89. (page 197) These addenda and supplements by Klein and Sommerfeld were added when Vol. IV appeared in 1910.

90. (page 197) Harold Crabtree (1874–1915) studied mathematics in Pembroke College, Cambridge. He subsequently became a teacher in the Charterhouse School, an elite English school for boys where he had formerly been a student. An obituary notice in the May 1916 volume of the *Mathematical Gazette* describes Crabtree's career at Charterhouse and the writing of his book on the top [Crabtree 1909] in the following words:

> A wrangler and a keen mathematician, he came to Charterhouse to teach Mathematics, but his deep piety leading him to claim a share in the teaching of Scripture, he was soon found to possess a remarkable knack in opening the eyes of boys to the beauties of Isaiah. Thus he gradually took up the teaching of English (his success with small boys in Shakespeare was remarkable) and, later on, of Classics in place of his mathematical work.
>
> In his mathematical teaching he was very lucid and particularly neat in the arrangement of his own work, so that he excelled in inculcating good style; he always had his desk full of papers and solutions neatly docketed and immediately available when he wanted, and he thoroughly understood how to work to a scheme.
>
> At the time that he published his delightful book on *Spinning Tops and Gyroscopic Motion* – which has obtained an international reputation – only a small percentage of his school work was mathematical, and he was always somewhat apologetic for having "accidentally" written such a book.

91. (page 197) Arthur Gordon Webster (1863–1923) studied physics in Berlin with Hermann von Helmholtz (1821–1894) and August Kundt (1839–1894). In 1890, Webster was hired as a physics instructor at Clark University in Worcester, Massachusetts, where Albert A. Michelson (1852-1931) was then the head of the department of physics. When Michelson moved to the University of Chicago in 1892, Webster became the department head at Clark, a position he held until his death in 1923. Webster's textbook on dynamics [Webster 1904] was highly regarded in both the United States and Europe. A biographical sketch of Webster in the *American Journal of Physics* [Duff 1938, pp. 185–186] quotes a review of Webster's book by Villino Volterra [sic; probably Vito Volterra (1860–1940)]:

> Webster had a great facility for learning and mastering all new ideas. His *Dynamics* is a work in which he has treated nearly all the branches of mechanics. A large number of theories are to be found there, condensed into a minimum number of pages, without losing either clarity or elegance or profundity. To describe it as a didactical work and as a manual useful to all physicists would be hardly adequate.

No record of the German edition by C. H. Müller (probably Conrad Heinrich Müller (1878–1953); Ph.D. Göttingen, 1904; privatdozent in Göttingen, 1908–1910; professor in the *Technische Hochschule* in Hannover, 1910–1948) has been found by the translators.

92. (page 197) August Otto Föppl (1854–1924) was professor of technical mechanics in the *Technische Hochschule* in Munich. One of his first doctoral students was the great aerodynamicist Ludwig Prandtl (1875–1953), who was invited by Felix Klein in 1904 to become the director of the new Institute for Technical Physics in the University of Göttingen. Prandtl later married the daughter of his teacher Föppl; the story of Prandtl's marriage is told in the autobiography of Prandtl's student Theodore von Kármán (1881–1963) [von Kármán 1967, p. 38]:

> Prandtl's personal life was particularly full of overtones of naïveté, which I became more and more familiar with over the years. One day, I recall, he came to the conclusion that he should get married. He was then nearing forty years of age. As a full professor, he had a permanent position at Göttingen, so he was secure financially. But he didn't know how to go about finding a wife. Finally he decided to write to Mrs. August Föppl, the wife of his mechanics teacher at Munich. The Föppls had two eligible daughters, one in her early twenties, the other in her late twenties. Prandtl asked for the hand of one of the daughters; he didn't specify which one.
>
> Flattered but puzzled, Mrs. Föppl called her daughters into a kind of emergency meeting and asked which of them wanted to marry Prandtl. Both daughters replied that Prandtl had been nice to them. He had made formal calls and had shown them funny mechanical toys, but had never spoken of love. There was a further family discussion, and it was decided to give Gertrude, the elder daughter, to

> Prandtl and to save the younger daughter for another candidate. Prandtl didn't object. The marriage turned out to be quite satisfactory, and the union produced two daughters.

Von Kármán was in Göttingen from 1906 to 1912. With allowance for his own vanity and embellishment due to the passage of time, von Kármán's autobiography gives an interesting glimpse of the personalities of Prandtl, Klein, and the other brilliant scientists and mathematicians in Göttingen at that time.

93. (page 198) Alfred George Greenhill (1847–1927) was a student of James Clerk Maxwell and a professor of mathematics in the Royal Military Academy at Woolwich, England. His work on the stereographic representation of the motion of the top is cited by Klein and Sommerfeld in Vol. II, Chapter IV, §1.

94. (page 198) This note on the Euler angles should probably refer to page 17.

95. (page 198) An English translation of the *Introductio in analysin infinitorum* was published in 1988 [Euler 1988]. The Euler angles appear in Chapter IV of the Appendix to Book II, as part of a discussion of coordinate changes for the representation of two-dimensional surfaces.

96. (page 198) Arthur Moritz Schoenflies (1853–1928) was a student of Karl Theodor Wilhelm Weierstrass (1815–1897) in the University of Berlin. In 1892 he was appointed to the newly created professorial chair of applied mathematics in Göttingen. He is best known for his application of group theory to crystallography [Schoenflies 1891]. The cited paper [Schoenflies 1910] is an elegant derivation of parametric representations for reflections, rotations, and screw motions. The decomposition of a rotation into two reflections is discussed in translators' note 20.

97. (page 198) In 1903, Felix Klein formed a *Vektorkommission* consisting of Arnold Sommerfeld, Ludwig Prandtl, and Rudolf Mehmke (1857–1944). Its goal was to standardize the definitions and notation of vector analysis. The cited paper by Mehmke, entitled *Vergleich zwischen der Vektoranalysis amerikanischer Richtung und derjenigen deutsch-italienischer Richtung* [Mehmke 1904], is a detailed comparison of the vector analysis of the American Josiah Willard Gibbs (1839-1903) with that of the German Hermann Günter Grassmann (1809–1877) and the Italian Giuseppe Peano (1858–1932). A major difference between

the two approaches lies in the concept of the bivector, which is used by Grassmann and Peano but not by Gibbs. The cited paper by Prandtl [Prandtl 1904] generalizes the American approach to the *physikalische Richtung* (the Gibbs analysis with the added distinction between polar vectors and axial vectors), and defines a third category that Prandtl calls the *trigonometrische Richtung* (the quaternion theory of Hamilton). Sommerfeld and Prandtl favored the physically oriented Gibbs analysis, while Mehmke advocated the more mathematical Grassmann–Peano analysis. Mehmke ends his paper with the following plea:

> Our considerations have shown that *the terminology of the German–Italian school is greatly preferable to that of Gibbs not only in logic and method, but also in practical respects.* And the terminology of Gibbs is the best of the numerous others that are used by the adherents of the American approach! Must not the diversification of terminology that is characteristic of this school and that is deplored by its adherents themselves, as opposed to the complete, unified conception and terminology of the German–Italian approach, be cause for general reflection?

Apparently not; the Gibbs vector analysis is now used almost universally for scientific work.

In 1908, Klein himself summarized the result of his commission [Klein 1908, pp. 157–158]:

> I would like to add a few words concerning the troublesome *question of notation in vector analysis.* There are, namely, a great many different symbols used for each vector operation, and it has unfortunately not yet been possible to produce a single generally accepted notation. Four years ago at the meeting of natural scientists at Kassel, we established a commission for this purpose; it members, however, could not completely agree with one another, and, since each had the good will to meet the others half-way, the sole result was that approximately three new notations were created! It appears to me, after this and similar experiences, that a real agreement in all such matters regarding spheres of common speech and notation is generally possible only if exceedingly important material interests stand behind it. Only under such pressure could the uniform electrical measurement system of the volt, ampère, and ohm be generally

> adopted in 1881 and then established by public legislation, since industry was in urgent need of such unity of calculation as the basis for all its designations. No such strong material interests stand as yet behind the vector calculus, and thus one must presently accept, for better or worse, that each individual mathematician will abide by his accustomed notation, which he feels the most convenient, or even—if he is somewhat dogmatically disposed—the only correct one.

98. (page 198) The use of six-vectors, which are now generally written as four-dimensional antisymmetric second-rank tensors, is discussed by Sommerfeld in his *Electrodynamics* [Sommerfeld 1964, §26].

99. (page 199) Hamilton's paper "On Quaternions" appeared in the Proceedings of the Royal Irish Academy in 1847 [Hamilton 1847]; the paper was read on November 11, 1844. In a footnote to the published paper, Hamilton makes the following comment, in his characteristic style, on Cayley's discovery of the rotation formula:

> The present writer desires to return his sincere acknowledgments to Mr. Cayley for the attention he has given to the Papers on Quaternions, published in the above-mentioned Magazine [Cayley 1845]: and gladly recognizes his priority, as respects the printing of the formula just now referred to. But while he conceives it to be very likely that Mr. Cayley, who had previously published in the Cambridge Mathematical Journal [Cayley 1843] some elegant researches on the rotation of bodies, may have perceived, not only independently, but at an earlier date than he did himself, the manner of applying quaternions to represent such a rotation; yet he hopes that he may be allowed to mention, that a formula differing only slightly in its notation from the formula (i) of the present abstract, with the corollaries there drawn respecting the composition of successive finite rotations, had been exhibited to his friend and brother Professor, the Rev. Charles Graves, of Trinity College, Dublin, in an early part of the month (October, 1844), which preceded that communication to the Academy, of which an account is given above.

100. (page 202) The title of the book by Hendrik Antoon Lorentz (1853–1928) is *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern* [Lorentz 1895]. A small excerpt of the book is included in the *The Principle of Relativity* [Lorentz 1952, pp. 3–7], an English translation of a collection of original papers on the theory of relativity.

101. (page 202) An English translation of Einstein's paper appears in the previously cited collection *The Principle of Relativity* [Lorentz 1952, pp. 35–65.].

102. (page 202) Poincaré's paper *"Sur la dynamique de l'électron* [Poincaré 1906] is available in his Œuvres [Poincaré 1954, pp. 494–550].

103. (page 202) The mathematician Hermann Minkowski (1864–1909) was born in Königsberg, where he was a schoolmate of Arnold Sommerfeld. Minkowski presented his space-time conception of relativity in a lecture to the *Kölner Naturforscher-Gesellschaft* in 1908. Sommerfeld attended the lecture, and provided notes for the transcription of the lecture that appears in the *The Principle of Relativity* [Lorentz 1952, pp. 73–91.].

104. (page 203) William Kingdon Clifford (1845–1879) studied mathematics in Trinity College, Cambridge, and became professor of mathematics and mechanics in University College, London. His work on the theory of quaternions is collected in his *Mathematical Papers* [Clifford 1882].

105. (page 203) Christian Hugo Eduard Study (1862–1930) taught mathematics in Leipzig, Marburg, and Johns Hopkins University before accepting a faculty position in Göttingen in 1894. A curiously titled doctoral dissertation on the life and work of Study was recently written in the Johannes Gutenberg-Universität in Mainz [Hartwich 2005].

106. (page 205) Eugen Stübler (1873–1930) was professor in the *Technische Hochschule* in Charlottenburg. *Der Impuls bei der Bewegung eines starren Körpers* [Stübler 1906] was his *Habilitationsschrift* in the *Technische Hochschule* in Stuttgart.

107. (page 205) In the cited *Encyklopädie* article [Stäckel 1908, pp. 581–589], Stäckel gives an interesting criticism of Euler's original derivation of his equations, and also discusses the subsequent derivations of Saint-Guilhem, Lagrange, and Poisson.

References

[D'Alembert 1990]. Jean D'Alembert, *Traité de dynamique*, Jacques Gabay, Sceaux, 1990; reprint of the original edition published by David, Paris, 1743; available at http://gallica.bnf.fr.

[Altmann 1989]. Simon L. Altmann, "Hamilton, Rodrigues, and the Quaternion Scandal," *Mathematics Magazine*, Vol. 62, 1989, pp. 291–308.

[Appell 1893]. Paul Émile Appell, *Traité de mécanique rationelle*, Gauthier-Villars, Paris, 1893.

[Arnold 1961]. Ronald N. Arnold and Leonard Maunder, *Gyrodynamics and Its Engineering Applications*, Academic Press, New York, 1961.

[Ball 1876]. Robert Ball, *The Theory of Screws*, Hodges and Porter, Dublin, 1876; available at http://gallica.bnf.fr.

[Bohnenberger 1818]. Johann Gottlob Friedrich Bohnenberger, *"Beschreibung einer Maschine, welche die Gesetze der Umdrehung der Erde um ihre Axe, und der Veränderung der Lage der Erdaxe zu erläutern dient," Annalen der Physik*, 1818, Bd. 60, pp. 60–71; available at http://gallica.bnf.fr.

[Boltzmann 1974]. Ludwig Boltzmann, "On the Indispensability of Atomism in Natural Science," in *Theoretical Physics and Philosophical Problems: Selected Writings*, D. Reidel, Boston, 1974, pp. 41–53.

[Brauer 1935]. Richard Brauer and Hermann Weyl, "Spinors in n dimensions," *American Journal of Mathematics*, Vol. 57, No. 2, 1935, pp. 425–449; available at http://www.jstor.org.

[Brocke 1980]. Bernhard vom Brocke, *Hochschul- und Wissenschaftspolitik in Preußen und im Deutsche Kaiserreich 1882–1907: das "System Althoff,"* Klett-Cotta, Stuttgart, 1980.

[Broelmann 2002]. Jobst Broelmann, *Intuition und Wissenschaft in der Kreiseltechnik, 1750 bis 1930*, Deutsches Museum, Munich, 2002.

[Burkhardt 1898]. Heinrich Burkhardt to Arnold Sommerfeld, 22 November 1898, Sommerfeld Papers Archive (NL 89, 006), Deutsches Museum, Munich.

[Cartan 1913]. Elie Cartan, "*Les groupes projectifs qui ne laissent invariante aucune mulitplicité plane*," *Bulletin de la Société Mathématique de France*, t. 41, 1913, pp. 53–96; available in Cartan, Elie, *Œuvres complètes*, Centre national de la recherche sciéntifique, Paris, 1984, ptie. 1, pp. 355–398.

[Cayley 1843]. Arthur Cayley, "On the Motion of Rotation of a Solid Body," *Cambridge Mathematical Journal*, t. III, 1843, pp. 224–232; available at http://www.hti.umich.edu/u/umhistmath in *The Collected Mathematical Papers of Arthur Cayley*, Vol. I, Cambridge University Press, 1889, pp. 28-35.

[Cayley 1845]. Arthur Cayley, "On certain results relating to Quaternions," *Philosophical Magazine*, t. XXVI, 1845, pp. 141–145; available at http://www.hti.umich.edu/u/umhistmath in *The Collected Mathematical Papers of Arthur Cayley*, Vol. I, Cambridge University Press, 1889, pp. 123–126.

[Cayley 1879]. Arthur Cayley, "On the correspondence of Homographies and Rotations," *Mathematische Annalen*, Bd. XV, 1879, pp. 238–240; available at http://gdz.sub.uni-goettingen.de.

[Clifford 1882]. William Kingdon Clifford, *Mathematical Papers*, Macmillan and Co., London, 1882; reprinted by Chelsea Publishing Company, 1968.

[Conway 2003]. John H. Conway and Derek A. Smith, *On Quaternions and Octonions: Their Geometry, Arithmetic and Symmetry*, A. K. Peters, Natick, 2003.

[Coriolis 1829]. Gaspard-Gustave de Coriolis, *Du calcul de l'effet des machines*, Carilian-Goeury, Paris, 1829; available at http://gallica.bnf.fr.

[Coriolis 1832]. Gaspard-Gustave de Coriolis, "*Sur le principe des forces vives dans les mouvemens relatifs des Machines*," *Journal de l'École Polytechnique*, t. 13, 1832, pp. 268–302.

[Coriolis 1835]. Gaspard-Gustave de Coriolis, "*Sur les équations du mouvement relatif des systèmes de corps*," *Journal de l'École Polytechnique*, t. 15, 1835, pp. 142–154.

[Crabtree 1909]. Harold Crabtree, *An elementary treatment of the theory of spinning tops and gyroscopic motion*, Longmans, Green and Co., London, 1909.

[Despeyrous 1884]. Théodore Despeyrous, *Cours de mécanique; avec des notes par m. G. Darboux*, A. Hermann, Paris, 1884.

[Diegel 1898]. Carl Diegel to Arnold Sommerfeld, 7 December 1898, reprinted in Michael Eckert and Karl Märker, eds., *Arnold Sommerfeld. Wissenschaftlicher Briefwechsel*, Band 1: 1892–1918, GNT-Verlag and Deutsches Museum, Diepholz, Berlin, Munich, 2000, pp. 97–101.

[Duff 1938]. A. Wilmer Duff, "Arthur Gordon Webster – Physicist, Mathematician, Linguist, and Orator," *American Journal of Physics*, Vol. 6, 1938, pp. 181–194; available at http://scitation.aip.org/ajp.

[Duhamel 1845]. Jean Marie Constant Duhamel, *Cours de mécanique de l'École polytechnique*, Bachelier, Paris, 1845–1846.

[Eckert 2003]. Michael Eckert, "The Practical Theorist: Sommerfeld at the Crossroads of Mathematics, Physics and Technology," *Philosophia Scientiæ*, Vol. 7, No. 2, 2003, pp. 165–188.

[Einstein 1926]. Albert Einstein, *"Die Ursache der Mäanderbildung der Flußläufe und des sogenannten Baerschen Gesetzes," Die Naturwissenschaften*, Vol. 14, 1926, pp. 223–224.

[Einstein 1982]. Albert Einstein, *Ideas and Opinions*, Three Rivers Press, 1982.

[Euler 1758]. Leonhard Euler, *"Du mouvement de rotation de corps solides autour d'un axe variable," Histoire de l'Académie Royale des Sciences et Belles-Lettres*, Berlin, 1758, pp. 154–193; available at http://bibliothek.bbaw.de.

[Euler 1988]. Leonhard Euler, *Introduction to Analysis of the Infinite* (English translation of *Introductio in analysin infinitorum*, Marc-Michel Bousquet, Lausanne, 1748), Springer-Verlag, New York, Book I, 1988, Book II, 1990.

[Fraser 1985a]. Craig Fraser, "D'Alembert's Principle: The Original Formulation and Applications in Jean's D'Alembert's *Traité de Dynamique* (1743)," *Centaurus*, Vol. 28, No. 1, 1985, pp. 31–61.

[Fraser 1985b]. Craig Fraser, "D'Alembert's Principle: The Original Formulation and Applications in Jean's D'Alembert's *Traité de Dynamique* (1743)," *Centaurus*, Vol. 28, No. 2, 1985, pp. 145–159.

[Gauss 1863]. Carl Friedrich Gauss, *Intensitas vis magneticae terrestris ad mensuram absolutam revocata*, in *Werke*, Band 5, pp. 81–120, Dieterich, Göttingen, 1863; available at http://gdz.sub.uni-goettingen.de.

[Gould 1973]. Douglas W. Gould, *The Top*, Clarkson N. Potter, New York, 1973.

[Hamilton 1847]. William Rowan Hamilton, "On Quaternions," *Proceedings of the Royal Irish Academy*, Vol. 3, 1847, pp. 1–16.

[Hamilton 1853]. William Rowan Hamilton, *Lectures on Quaternions*, Hodges and Smith, Dublin, 1853.

[Hartwich 2005]. Yvonne Hartwich, *"Eduard Study (1862–1930)—ein mathematischer Mephistopheles im geometrischen Gärtchen,"* doctoral dissertation in the Johannes Gutenberg-Universität, Mainz, 2005; available at http://deposit.ddb.de.

[Hayward 1856]. Robert Baldwin Hayward, "On a Direct Method of estimating Velocities, Accelerations, and all similar Quantities with respect to Axes moveable in any manner in Space, with Applications," *Transactions of the Cambridge Philosophical Society*, Vol. 10, Part I, 1856, pp. 1–20.

[Hermite 1885]. Charles Hermite, *Sur quelques applications des fonctions elliptiques*, Gauthier-Villars, Paris, 1885; available at http://historical.library.cornell.edu/math.

[Hermite 1905]. Charles Hermite, *Œuvres de Charles Hermite*, t. III, pp. 266–418, Gauthier-Villars, Paris, 1905–1917; available at http://-www.hti.umich.edu/u/umhistmath.

[Herschel 1851]. John Frederick William Herschel, *Outlines of Astronomy*, Longman, Brown, Green and Longmans, London, 1851; available at http://gallica.bnf.fr.

[Hertz 1894]. Heinrich Hertz, *Die Prinzipien der Mechanik*, in *Gesammelte Werke*, Band III, J. A. Barth, Leipzig, 1894; available at http://-gallica.bnf.fr.

[Hess 1882]. Wilhelm Hess, "*Ueber das Gyroskop,*" *Mathematische Annalen*, Bd. XIX, 1882, pp. 121–154; available at http://gdz.sub.uni-goettingen.de.

[Hess 1887]. Wilhelm Hess, "*Ueber das Gyroskop bei allgemeinster Wahl des zur Bewegung anregenden Momentankräftesystems,*" *Mathematische Annalen*, Bd. XXIX, 1887, pp. 500–580; available at http://-gdz.sub.uni-goettingen.de.

[Jacobi 1850]. Carl Gustav Jacob Jacobi, "*Sur la rotation d'un corps*," *Journal für die reine und angewandte Mathematik*, Bd. 39, 1850, pp. 293–350; available at http://gdz.sub.uni-goettingen.de.

[Jacobs 1977]. Konrad Jacobs, ed., *Felix Klein. Handschriftlicher Nachlaß*, Universität Erlangen, Erlangen, 1977.

[Jouffret 1874]. Esprit Pascal Jouffret, *"Théorie élémentaire du mouvement du gyroscope, de la toupie, et du projectile oblong," Revue d'artillerie*, Vol. 4, 1874, pp. 245–264, 547–558.

[Jouffret 1903]. Esprit Pascal Jouffret, *Traité élémentaire de géométrie à quatre dimensions et introduction à la géométrie à n dimensions*, Gauthier-Villars, Paris, 1903.

[von Kármán 1967]. Theodore von Kármán, *The Wind and Beyond*, Little, Brown and Company, Boston, 1967.

[Kirchhoff 1876]. Gustav Robert Kirchhoff, *Vorlesungen über mathematische Physik*, Bd. I, *Mechanik*, B. G. Teubner, Leipzig, 1876.

[Klein 1875]. Felix Klein, "*Ueber binare Formen mit linearen Transformationen in sich selbst*," *Mathematische Annalen*, Bd. IX, 1875, pp. 183–208; available at http://gdz.sub.uni-goettingen.de.

[Klein 1877]. Felix Klein, "*Weitere Untersuchungen über das Ikosaeder*," *Mathematische Annalen*, Bd. XII, 1877, pp. 503–560; available at http://gdz.sub.uni-goettingen.de.

[Klein 1884]. Felix Klein, *Vorlesungen über das Ikosaeder und die Auflösung der Gleichungen vom fünften Graden*, B. G. Teubner, Leipzig, 1884; available at http://gallica.bnf.fr; 1913 English translation by George Gavin Morrice republished by Dover Publications, New York, 1956.

[Klein 1896a]. Felix Klein, *Ausgewählte Kapitel der Zahlentheorie*, Vol. 1, (Lecture held in winter semester 1895/96, elaborated by A. Sommerfeld), B. G. Teubner, Leipzig, 1896.

[Klein 1896b]. Felix Klein, *Ausgewählte Kapitel der Zahlentheorie*, Vol. 2 (Lecture held in summer semester 1896, elaborated by A. Sommerfeld and Philip Furtwängler), B. G. Teubner, Leipzig, 1896.

[Klein 1897]. Felix Klein, *The Mathematical Theory of the Top*, Lectures delivered on the occasion of the sesquicentennial celebration of Princeton University, Charles Scribner's Sons, New York, 1897; reprinted by Dover Publications, 2004.

[Klein 1899]. Felix Klein to Arnold Sommerfeld, 15 November 1899, Deutsches Museum Archive (HS 1977-28/A, 170), Munich.

[Klein 1907]. Felix Klein to Arnold Sommerfeld, 20 November 1907. Deutsches Museum Archive (HS 1977-28/A, 170), Munich.

[Klein 1908]. Felix Klein, *Elementarmathematik vom höheren Standpunkte aus. Teil I: Arithmetik, Algebra, Analysis*, B. G. Teubner, Leipzig, 1908; available at http://books.google.com; English translation *Elementary Mathematics from an Advanced Standpoint. Arithmetic, Algebra, Analysis*, by E. R. Hedrick and C. A. Noble, Dover, New York, 2004.

[Klein 1922]. Felix Klein, *Gesammelte mathematische Abhandlungen*, Vol. 2, *Anschauliche Geometrie, Substitutionsgruppen und Gleichungstheorie, zur mathematischen Physik*, edited by R. Fricke and H. Vermeil, Springer, Berlin, 1922; 1973 reprint by Springer-Verlag available at http://gallica.bnf.fr.

[Lagrange 1788]. Joseph-Louis Lagrange, *Méchanique Analitique*, Veuve Desaint, Paris, 1788; available at http://gallica.bnf.fr.

[Lamb 1945]. Horace Lamb, *Hydrodynamics*, Dover Publications, New York, 1945.

[Liebmann 1898]. Heinrich Liebmann, "*Classification der Kreiselprobleme nach der Art der zugehörigen Parametergruppe*," *Mathematische Annalen*, Bd. 50, 1898, pp. 51–67; available at http://gdz.sub.uni-goettingen.de.

[Lorentz 1895]. Hendrik Lorentz, *Versuch einer Theorie der electrischen und optischen Erscheinungen in bewegten Körpern,* E. J. Brill, Leiden, 1895; available at http://www.historyofscience.nl.

[Lorentz 1952]. Hendrik Lorentz et al., *The Principle of Relativity*, Dover Publications, New York, 1952.

[Mach 1883]. Ernst Mach, *Die Mechanik in ihrer Entwickelung historisch-kritisch dargestellt*, F. A. Brockhaus, Leipzig, 1883.

[Maxwell 1873]. James Clerk Maxwell, *A Treatise on Electricity and Magnetism*, Clarendon Press, Oxford, 1873; available at http://gallica.bnf.fr.

[Maxwell 1965]. James Clerk Maxwell, *The Scientific Papers of James Clerk Maxwell*, Dover Publications, New York, 1965.

[Maxwell 1991]. James Clerk Maxwell, *Matter and Motion*, Dover Publications, New York, 1991.

[Mehmke 1904]. Rudolf Mehmke, *"Vergleich zwischen der Vektoranalysis amerikanischer Richtung und derjenigen deutsch-italienischer Richtung," Jahresbericht der deutschen Mathematiker-Vereinigung*, Bd. 13, pp. 217-228, 1904; available at http://gdz.sub.uni-goettingen.de.

[Moebius 1885]. August Moebius, *Gesammelte Werke*, Bd. 3, S. Hirzel, Leipzig, 1885–87; available at http://www.hti.umich.edu/u/umhistmath.

[Newton 1962]. Isaac Newton, *Mathematical Principles of Natural Philosophy*, Andrew Motte translation revised by Florian Cajori, University of California Press, Berkeley and Los Angeles, 1962.

[Perry 1890]. John Perry, *Spinning tops; the "Operatives' lecture" of the British Association meeting at Leeds, 6th September 1890*, Society for Promoting Christian Knowledge, London, E. S. Gorham, New York, 1910.

[Poinsot 1811]. Louis Poinsot, *Éléments de statique*, Volland, Paris, 1811; available at http://gallica.bnf.fr.

[Poinsot 1834]. Louis Poinsot, *Théorie nouvelle de la rotation des corps*, Bachelier, Paris, 1834. English translation: *Outlines of a New Theory of Rotary Motion*, Charles Whitley, R. Newby, Cambridge, 1834; available at http://kmoddl.library.cornell.edu. German translation: *Neue Theorie der Drehung der Körper*, Karl Heinrich Schellbach, Hayn, Berlin, 1851.

[Poinsot 1851]. Louis Poinsot, *"Théorie nouvelle de la rotation des corps," Journal de mathématiques pures et appliquées*, t. 16, 1851, pp. 9–129, 289–336; available at http://gallica.bnf.fr.

[Poincaré 1906]. Henri Poincaré, *"Sur la dynamique de l'électron," Rendiconti del Circolo Matematico di Palermo*, t. XXI, pp. 129–175, 1906.

[Poincaré 1954]. Henri Poincaré, *Œuvres de Henri Poincaré*, t. IX, Gauthier-Villars, Paris, 1954.

[Prandtl 1904]. Ludwig Prandtl, *"Über die phsysikalische Richtung in der Vektoranalysis," Jahresbericht der deutschen Mathematiker-Vereinigung*, Bd. 13, pp. 436–449, 1904; available at http://gdz.sub.uni-goettingen.de.

[Résal 1873] Henri Amé Résal, *Traité de mécanique générale: comprenant les leçons professées à l'École polytechnique*, Gauthier-Villars, Paris, 1873–1881.

[Robbin 2006]. Tony Robbin, *Shadows of Reality*, Yale University Press, 2006.

[Rodrigues 1840]. Benjamin Olinde Rodrigues, *"Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire," Journal de Mathématiques Pures et Appliquées*, t. 5, 1840, pp. 380–440; available at http://gallica.bnf.fr.

[Routh 1960]. Edward John Routh, *Dynamics of a system of rigid bodies*, Part I, Dover Publications, New York, 1960.

[Rowe 1989]. David E. Rowe, "Klein, Hilbert, and the Göttingen Mathematical Tradition," *Osiris*, 2nd Series, Vol. 5, "Science in Germany: The Intersection of Institutional and Intellectual Issues," 1989, pp. 186–213.

[de Saint-Germain 1887] Albert Leon de Saint-Germain, *Résumé de la théorie du mouvement d'un solide autour d'un point fixe, à l'usage des candidats à la licence és-sciences*, Extrait du Bulletin mensuel des facultés de Caen, Paris, 1887.

[Scheffler 1866]. Hermann Scheffler, "*Imaginäre Arbeit, eine Wirkung der Centrifugal- und Gyralkraft*," *Zeitschrift für Mathematik und Physik*, Bd. 11, 1866, pp. 93–151.

[Schell 1879]. Wilhelm Schell, *Theorie der Bewegung und der Kräfte*, Bd. I, B. G. Teubner, Leipzig, 1879.

[Schoenflies 1891]. Arthur Schoenflies, *Krystallsysteme und Krystallstructur*, B. G. Teubner, Leipzig, 1891; reprinted by Springer-Verlag, Berlin, 1984.

[Schoenflies 1910]. Arthur Schoenflies, *Einfache Ableitung der Parameterformeln für Bewegungen und Umlegungen,*" *Rendiconti del Circolo matematico di Palermo,* t. XXIX, 1910, pp. 329–339.

[Sommerfeld 1892]. Arnold Sommerfeld, *"Mechanische Darstellung der electromagnetischen Erscheinungen in ruhenden Köpern"*, *Annalen der Physik*, Bd. 46, pp. 139–151, 1892; available in Arnold Sommerfeld, *Gesammelte Schriften*, Bd. I, pp. 404–416, Friedr. Vieweg & Sohn, Braunschweig, 1968.

[Sommerfeld 1896]. Arnold Sommerfeld, *"Mathematische Theorie der Diffraction"*, *Mathematische Annalen*, Bd. 47, 1896, 317–347; English translation *Mathematical Theory of Diffraction*, Birkhäuser, Boston, 2004.

[Sommerfeld 1897a]. Arnold Sommerfeld to Felix Klein, 18 March 1897, reprinted in Michael Eckert and Karl Märker, eds., *Arnold Sommerfeld. Wissenschaftlicher Briefwechsel,* Band 1: 1892–1918, GNT-Verlag and Deutsches Museum, Berlin, Diepholz, Munich, 2000, pp. 71–72.

[Sommerfeld 1897b]. Arnold Sommerfeld to Johanna Höpfner, 16 January 1897, private collection.

[Sommerfeld 1898a]. Arnold Sommerfeld to Felix Klein, 15 May 1898, Göttingen, Staats- und Universitätsbibliothek, Klein Papers (Klein 11, 1045).

[Sommerfeld 1898b]. Arnold Sommerfeld to Felix Klein, 15 December 1898, Göttingen, Staats- und Universitätsbibliothek, Klein Papers (Klein 11, 1051).

[Sommerfeld 1899]. Arnold Sommerfeld to Felix Klein, 2 November 1899, Göttingen, Staats- und Universitätsbibliothek, Klein Papers (Klein 11, 1055).

[Sommerfeld 1900]. Arnold Sommerfeld to Felix Klein, 8 November 1900, Göttingen, Staats- und Universittsbibliothek, Klein Papers (Klein 11, 1061); reprinted in Michael Eckert and Karl Märker, eds., *Arnold Sommerfeld. Wissenschaftlicher Briefwechsel*, Band 1: 1892–1918, GNT-Verlag and Deutsches Museum, Berlin, Diepholz, Munich, 2000, pp. 182-184.

[Sommerfeld 1964]. Arnold Sommerfeld, *Electrodynamics*, Academic Press, New York, 1964.

[Stäckel 1908]. Paul Gustav Stäckel, "*Elementare Dynamik der Punktsysteme und starren Körper*," *Encyklopädie der mathematische Wissenschaften*, Bd. IV, 1 Teilband, 1901–1908, pp. 443–684; available at http://gdz.sub.uni-goettingen.de.

[Stübler 1906]. Eugen Stübler, *Der Impuls bei der Bewegung eines starren Körpers,* B. G. Teubner, Leipzig, 1906.

[Sylvester 1866]. James Joseph Sylvester, "On the Motion of a Rigid Body acted on by no external Forces," *Philosophical Transactions of the Royal Society of London*, Vol. 156, 1866, pp. 757–779; available at http://www.jstor.org.

[Synge 1949]. John L. Synge and Byron A. Griffith, *Principles of Mechanics*, McGraw-Hill, 1949.

[Tait 1866]. Peter Guthrie Tait, "Sir William Rowan Hamilton," *North British Review*, Vol. 45, 1866, pp. 37–74.

[Tait 1898]. Peter Guthrie Tait, *Scientific Papers*, Cambridge University Press, 1898; available at http://gallica.bnf.fr.

[Thomson 1879a]. William Thomson and Peter Guthrie Tait, *Elements of Natural Philosophy*, Cambridge University Press, Cambridge, 1879.

[Thomson 1879b]. William Thomson and Peter Guthrie Tait, *Treatise on Natural Philosophy*, Vol. I, Part I, Cambridge University Press, Cambridge, 1879.

[Thomson 1883]. William Thomson and Peter Guthrie Tait, *Treatise on Natural Philosophy*, Vol. I, Part II, Cambridge University Press, Cambridge, 1883.

[Thomson 1890]. William Thomson, *Mathematical and Physical Papers*, Vol. III, C. J. Clay and Sons, London, 1890; available at http://-gallica.bnf.fr.

[Tomonaga 1997]. Sin-itiro Tomonaga, *The Story of Spin*, University of Chicago Press, Chicago, 1997.

[Webster 1904]. Arthur Gordon Webster, *The dynamics of particles and of rigid, elastic, and fluid bodies*, B. G. Teubner, Leipzig, 1904; available at http://books.google.com; reprinted by Dover Publications, New York, 1959;

[Weinhold 1881]. Adolf Ferdinand Weinhold, *Physikalische Demonstrationen*, Quandt & Händel, Leipzig, 1881.

[Winkelmann 1904]. Max Winkelmann, *Zur Theorie des Maxwell'schen Kreisels*, Inaugural Dissertation, Göttingen, 1904; available at http://-books.google.com.

Index